DISCONTINUOUS RECEPTION AND TRANSMISSION (DRX/DTX) STRATEGIES IN LONG TERM EVOLUTION (LTE) FOR VOICE-OVER-IP (VOIP) TRAFFIC UNDER BOTH FULL-DYNAMIC AND SEMI-PERSISTENT PACKET SCHEDULING POLICIES

by

DARIO VINELLA MICHELE POLIGNANO

Project Group: 996

A Thesis Report submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Department of Electronic Systems
Radio Access Technology Section
Aalborg University

Supervisors:
Prof. Troels B. Sørensen (Aalborg University)
Ph.D. Jeroen Wigard (Nokia Siemens Networks)

November 20th, 2009

Number of copies: 5

Number of pages: 123

Abstract

New generation mobile communication systems like Long Term Evolution (LTE) aim to deploy to customers a new mobile experience providing higher data rates and lower latencies that can make wireless devices a great platform to run a new whole set of services and applications, that was unimaginable just some years ago. However, the energy demands of battery powered devices, needed to support these new software applications substantially exceeds the capacity of the actual battery technology. The development of new architectures and procedures to build power-efficient and power-aware systems has become one of the main purposes in the design of new generation wireless networks.

LTE exploits the idea of Discontinuous Reception (DRX) and Discontinuous Transmission (DTX), to provide a concrete solution to the power saving issue. The main point of this functionality makes the terminal to not continuously monitor control channels, allowing it to turn the radio frequency modem in sleep state for long periods, activating it only in well defined, suitable, instants.

The objective of this thesis is to analyze the transmission of VoIP traffic over LTE networks trying to find the best solution to efficiently deliver data, fitting the DRX/DTX functionality to provide power saving and not perceptibly degrade the user experience.

VoIP traffic has been studied with a bidirectional model, that emulates the interaction of speakers in a phone call. Two suitable scheduling policies for this kind of traffic, semi-persistent (SPS) and dynamic (DS), have been investigated underlining pros and cons of each solution. It has also introduced a simple control for SPS that can improve its performances. Other modeled entities include CQI reporting module, link adaptation algorithms, transmission and retransmission managers, a detailed implementation of a DRX/DTX framework and a simple solution to track power consumption statistics. The study involves the use of different configuration for channel quality reporting, link adaptation, DRX/DTX behavior with the extension of short DRX/DTX too. Every combination has been tested in single user, and multiple users scenario with 250 users all positioned at cell edge.

The use of DRX/DTX functionality with SPS has been proved to provide good power saving effects due to the general synchronization between data source, SPS and DRX/DTX behaviors. This configuration generally induces acceptable delays in single user scenarios, but multiple users scenarios with too long DRX/DTX cycles can generate losses of about 4% of packets, and delays that propagate for the whole talk-spurt. Dynamic scheduling has shown to make good use of its channel quality tracking, performing an instantaneous correct estimation of needed resources. The cost of dynamic scheduling resides in its more frequent signaling activity, that impacts to a 33% higher power consumption compared to the SPS case. The common outcome of SPS and DS analysis is that the short DRX/DTX functionality does not bring further power savings to any of the studied scheduling solutions, due to VoIP traffic pattern.

Preface

This thesis represents the conclusion work of a project developed at the Radio Access Technology (RATE) section of Aalborg University in collaboration with Nokia Siemens Networks research center of Aalborg, under the supervision of Associate Professor Troel B. Sørensen (Aalborg University, Denmark) and Dr. Jeroen Wigard (Nokia Siemens Networks, Aalborg, Denmark). The project has been co-financed by Nokia Siemens Networks, Aalborg and Erasmus Lifelong Learning Programme.

There are several people without whom this thesis would not have been at all possible and whom we need to thank. Our first thanks go to Prof. Gennaro Boggia and Prof. Luigi Alfredo Grieco that gave us the chance and contacts to develop our master thesis at the Nokia Siemens Networks research center of Aalborg.

A grateful acknowledgment goes to all the people we met at NSN and university, that have contributed to the completion of this project. We want to express our gratitude to Dr. Jeroen Wigard that directed, advised and supported our daily work. His help has always been forthcoming for every question or trouble, steering us in the right direction, but allowing us to make this thesis our own work.

We would like to thank Prof. Troel B. Sørensen for the patient revision of all our work and for all suggestions and advices he gave us in numerous constructive discussions.

Many thanks as well to all the generous support of all NSN employee and students, in particular Daniela, Claudio C. and Claudio R. always kind to give us useful tips and reply to our persistent questions.

Contents

List of Abbreviations

3G	Third Generation
3GPP	3rd Generation Partnership Project
AMR	Adaptive Multirate
AVI	Actual Value Interface
BLER	Block Error Rate
BSR	Buffer Status Report
CC	Chase Combining
CP	Cycling Prefix
CQI	Channel Quality Indicator
CS	Circuit Switched
DFT	Discrete Fourier Transform
DFT-SOFDM	DFT Spread OFDM
DVB-H	Digital Video Broadcasting - Handheld
DVB-T	Digital Video Broadcasting - Terrestrial
EDGE	Enhanced Data rates for GSM Evolution
EESM	Exponential Effective SINR Metric
eNodeB	E-UTRAN Node B
FDMA	Frequency Division Multiple Access
FDPS	Frequency domain packet scheduler
GBR	Guarantee Bit Rate
GERAN	GSM EDGE Radio Access Network(GERAN)
GSM	Global System for Mobile Communications
HARQ	Hybrid Automatic Repeat Request
HSPA	High Speed Data Access

IDFT	Inverse Discrete Fourier Transform
IMT-2000	International Mobile Telecommunications-2000
IP	Internet Protocol
IR	Incremental Redundancy
ISI	Inter-symbol Interference
LTE	Long Term Evolution
MBSFN	Multicast Broadcast Single Frequency Network
MCS	Modulation and Coding Scheme
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
MME	Mobility Management Entity
MRC	Maximum Ratio Combining
NAS	Non-Access Stratum
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PAR	Peak To Average Ratio
PDCCH	Physical Downlink Control Channel
PDSCH	Physical Downlink Shared Channel
PRB	Physical Resource Block
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel
QAM	Quadrature Amplitude Modulation
RAN	Radio Access Network
RAT	Radio Access Technology
RB	Resource Block
RF	Radio Frequency
ROHC	Robust Header Compression
RR	Resource Request
RRM	Radio Resource Management
SAE	System Architecture Evolution

SAW	Stop and wait
SC	Single Carrier
SC-FDMA	Single Carrier Frequency Division Multiple Access
SID	Silence Descriptor
SIMO	Single Input Multiple
SINR	Signal to Interference-Plus-Noise Ratio
SR	Scheduling Request
SRS	Sounding Reference Signal
TSG	Technical Specifications Group
TTI	Transmission Time Interval
VAD	Voice Activity Detector
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network

Chapter 1

Introduction

The availability of more and more powerful hardware in mobile computation systems makes these devices the future platform for the development of a large set of new applications that will open new business models and will increase the market penetration in current society. However, the multitude and complexity of components needed to implement a large number of communications protocols makes the evaluation of power efficiency one of the most tight issues. In fact, the performance and utility of distributed systems is impacted by the availability of participating nodes. Each one of these must be connected to another to ensure remote task execution and data consistency, in an *always on* manner. This particular objective, hard to be achieved in wired network, is even more tricky in mobile environments. Mobile devices show the unwelcome peculiarity of a finite lifetime [29]. The power consumption of PDAs and smartphones represents a very complex problem, because of the number of physical components of the system, each with its own consumption value (backlight, bluetooth module, Wireless Local Area Network (WLAN) module, phone module, other circuits, etc.). Nevertheless, common behavior of physical components is strongly related to the executed operations type, so the software application layer has a big influence on the energy consumption too [30].

The important fact is that the energy demands of battery powered devices, added to the software applications behaviors substantially exceed the capacity of the actual battery technology. The development of new architectures and procedures to build power-efficient and power-aware systems has become one of the main purposes in the design of new mobile communication systems.

New generation mobile communication systems like Long Term Evolution (LTE) aim to deploy to customers a new mobile experience providing higher data rates and lower latencies that can make wireless devices a great platform to run a new whole set of services and applications, that was unimaginable just some years ago. One of the key objectives to support this evolution is the improved efficiency of power utilization: LTE exploits the idea of Discontinuous Reception (DRX) and Discontinuous Transmission (DTX), described in section 2.5, to achieve it. The main point of this functionality makes the terminal to not monitor the control channels continuously, but only in well defined instants, turning the RF modem in a sleep state as much as possible. In fact, control data are low-frequency and low bit-rate, so continuously listening the related channels represents a big waste of power resource.

A favorable coordination between the behavior of data transmission and DRX/DTX management can lead to high power saving values: this target can be reached in an easier way when control data and payload data flows can be finely predicted i.e. when the signal data source is well known, like Voice over Internet Protocol (VoIP) traffic.

The VoIP traffic represents a growing portion of the whole flow conveyed every day on packet networks. It is made of periodical transmission of small packets, and lend itself to a fine cooperation with DRX/DTX functionality.

In the following section we will define the current status of VoIP technology, its growing magnitude in mobile services, and its satisfaction metrics.

Voice on packet networks

Voice over Internet Protocol (VoIP) is an emerging service that is progressively increasing market quota over traditional telephony. IP based telephony is moving fast from a low scale phenomenon to large scale carrier service. The main advantages of this evolution for providers are less expensive and more scalable networks that can deliver a good quality service, integration of voice and data applications and lower bandwidth requirements. VoIP service increases its numbers more and more every year: the estimated number of VoIP users for 2011 will approximately be 250 millions [9].

VoIP moves telephony digital data from circuit switched networks, to packet switched ones since they can deliver data bits cost-effectively, making space for additional innovative services. Recently, the VoIP interest has approached cellular networks too: as predicted few years ago, the same evolution and switch from circuit to packet switched networks is happening in the mobile environment. A lot of current studies aim to find the best solutions to provide high quality and satisfying user-experience for mobile VoIP customers.

The main service requirements for VoIP are described by its satisfaction metrics: quality, coverage and capacity. 3GPP Rel.99 specifications show that these objectives are at least as good as Circuit Switched (CS) speech [3].

The fundamental quality parameter for the evaluation of the user experience in interactive talks is the perceived delay.

The total amount of transmission time over networks with digital segments is mainly composed by processing and propagation delays. The design of telecommunications networks has to take into account that each delay contributes to the total one experienced by user, because higher delays mean really unsatisfying interactive experiences.

The historical value of 400ms has been widely used in networks oriented to voice telephony services, indeed transmission time is the most valuable factor that impacts over user or terminal interactivity [15]. ITU-T recommends the following limits for one-way transmission time

1. 0 to 150ms: acceptable for most users applications
2. 150 to 400ms: acceptable provided that operators are aware that the transmission delay can impact over the service or application quality
3. above 400ms: unacceptable for almost all voice service networks

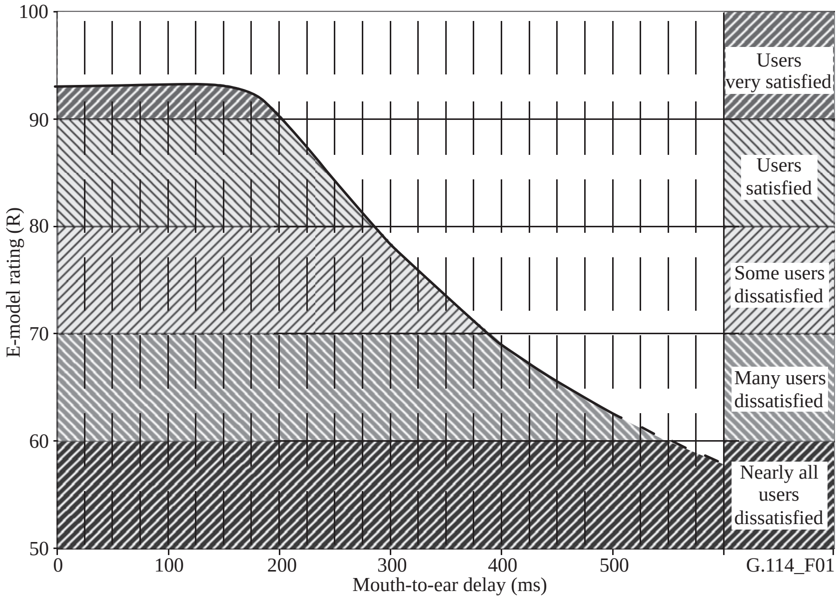


Figure 1.1: VoIP satisfaction thresholds [16]

The whole delay budget allocation takes into account the propagation time determined by distance and speed of signal in transmission facility, geographical distance between interfaces, routing choices for circuit switched networks.

VoIP system serves voice data over packet switched networks, so the delay budget should be divided in a different way. Moreover, current IP networks can only offer best-effort services and allow new traffic to enter even beyond capacity limits, causing both packet losses and significant delays [9].

The delays introduced by a VoIP network system are generated by data plane – that includes compression, framing, coding, transmission, etc. – and control plane – services for resources provisioning, admission control, resource reservation, connection management.

1.1 Thesis objectives

The study described in this thesis investigates the best parametrization of DRX/DTX functionality to achieve the best trade off between VoIP user experience and battery efficiency. The use of a conversational speech model makes possible to finely emulate a bidirectional call, showing all the behavioural patterns of two speakers. Since the transmission of a VoIP call can radically change according to the way the packets are scheduled, we also analyze the differences between two scheduling policies, adapting the DRX/DTX concepts to both of them.

Problem formulation

On the basis of two different scheduling policies, Semi-Persistent and Dynamic resource allocation for VoIP traffic, we will reply to the following questions in the coming chapters:

- Which performance Semi-Persistent and Dynamic scheduling policies show with VoIP traffic?
- Can Semi-Persistent Packet Scheduling guarantee low error rates in packets transmission, to be profitably used for VoIP data scheduling?
- Which link adaptation techniques can be implemented to keep low error rates in both Semi-Persistent and Dynamic scheduling and which parametrization should be utilized?
- How DRX/DTX functionality impacts to packet delivery delay?
- How DRX/DTX functionality impacts to power consumption with VoIP traffic?

1.2 Study Methodology

To fulfill the objectives of the thesis, we have first of all analyzed the current status of LTE standardization at 3GPP website (http://www.3gpp.org) and a set of technical papers and reports provided by Nokia Siemens Networks. We have then implemented the needed logics into a system level simulator to perform a large set of simulations and analysis over a wide range of parameters setting. The useful results obtained with extensive campaigns have been then processed and post-processed to find the average behavior of the system.

1.2.1 The simulator

All the studies conducted in this report are performed with the help of a MATLAB®system level simulator expressly developed for the purpose. The adopted simulation methodology is known as *Monte Carlo statistical method*, a technique used to simulate physical and mathematical systems when the analyzed system behavior is influenced by a large set of parameters and random actions, with the result of an huge amount of permutations of possible behaviors. In these cases a mathematical analysis of the system is unfeasible. Basically, the Monte Carlo method consists in a large number of simulations for each set of parameters, to be able to extract a mean of the results, that should track the probability distribution of the system behavior under examination.

The simulator implements several modules that emulate the behavior of real entities involved in both downlink and uplink transmissions in a cellular system, and especially in LTE. The crucial systems and features are implemented in details to accomplish the specifications fixed by the main LTE specification authority, the 3GPP. However, not all the functionalities of LTE have been finely standardized and tuned, so we have proposed some assumptions mainly modeled on Nokia Siemens Networks reports and technical papers. Finally, some of the marginal system characteristics that do not significantly impact to the performed analysis have been subject of simplifications or rough modeling.

The link level data needed to perform the analysis have been provided by two Nokia Siemens Networks simulators, and will be described in the following chapters.

The most important objectives of the simulator implementation have been a detailed modeling of packet scheduling and DRX/DTX control entities, as well as a good voice model. The innovative part is indeed based on the particular conversational model implemented: the availability of bidirectional call traces allows to perform more accurate analysis of VoIP calls and especially on DRX/DTX behavior, that, as it will be shown in next chapters, is strictly influenced by the combined action of uplink and downlink transmissions.

1.3 Main contributions

The main contribution of this master thesis project has been the implementation of a MATLAB®tool able to emulate the dynamic of two speakers engaged in a bidirectional call and all the LTE features involved in this scenario.

For this reason, the first step to realize our objectives has required the introduction of a new voice model that can provide call traces able to emulate real interaction phenomena between two speakers and take into account the correlation of the their activity, instead of the 50%/50% voice activity model, mostly used in available researches.

The successive steps have lead to the complete realization of the simulator with all the LTE functionalities needed for the realization of this study, including a new simple open loop control for the semi-persistent scheduling that seems to provide conservative but not wasteful allocation.

The main outcome of the following analysis is the impact of the DRX/DTX features on SPS and DS performance and the gain it can provide on power saving. Another important conclusion concerns the utilization of short DRX/DTX that seem to be not useful with VoIP traffic. More details will be described in the following chapters.

1.4 Thesis outline

The following chapters of the thesis are organized in this way:

- Chapter 2 provides a description of LTE system, the main purposes of the project, the actual specification status and implementation. The main functionalities and characteristics closely involved in the thesis analysis, like radio resource management and DRX/DTX are explained in detail.
- Chapter 3 presents the actual modeling and implementation of LTE features. The chapters also describes the conversational speech model used in the simulator, with a general overview of the voice codec.

- Chapters 4 and 5 enclose the assumptions, implications and analysis of the use of semi-persistent and dynamic packet schedulers for VoIP traffic in LTE. The assumed scenarios are described and the relative results explained, and compared showing the main satisfaction and performance criteria for this kind of review.

- Chapter 6 performs a close comparison of the two scheduling policies giving an overall idea of the main benefits and losses of each analyzed configuration.

- Chapter 7 summarizes the report conclusions indicating possible future works.

The reader with previous knowledge of LTE system can have a look of the main contributions and outcomes beginning from chapter 3.

Chapter 2

LTE

The pervasiveness of the present-day broadband cabled services has made the Internet generation used to high-speed connections and to a whole new set of services usually available exclusively from fixed locations. However, mobile broadband is currently becoming a reality with third generation (3G) technologies: out of estimated 1.8 billion people who will be served by broadband by 2012, two-thirds will be mobile broadband customers [11].

During the Radio Access Networks (RAN) Future Workshop, it has become clear that 3G long-term evolution has to meet this future services demand, with the constraint to remain competitive for a long time frame i.e. coming decades. A consistent amount of new technologies, like Orthogonal Frequency Division Multiplexing (OFDM) with flexible and broader Radio Frequency (RF), linked with enhancements in both Universal Terrestrial Radio Access Network (UTRAN) architecture and UTRA radio interface were presented as candidates to fulfill this purpose.

The main justifications for this evolution path reside in the constant increase of customer utilization of new and traffic-consuming services: people can already successfully browse the Internet, manage emails, send and receive audio-visual contents with High Speed Data Access (HSPA) enabled devices, like smartphones, PDAs, notebook, dongles, but more and more services with higher data rate, low latency and availability are approaching the mobile world.

The starting point of Long Term Evolution (LTE) was fixed when 3rd Generation Partnership Project (3GPP)[1], in the 3GPP Technical Specifications Group (TSG) RAN #26 meeting, approved the study item description on “Evolved UTRA and UTRAN” [1].

LTE performances and technologies have been evaluated with a number of so-called checkpoints and the final results agreed on during 3GPP plenary sessions in May 2007 in South Korea. Some of the focused point were:

- Significantly increased peak data rate e.g. 100 Mbps (downlink) and 50 Mbps (uplink)
- Increase “cell edge bit rate” whilst maintaining same site locations as deployed today

[1]3GPP is a collaboration agreement founded in 1998 that gathers some telecommunications standards bodies “Organizational Partners” like ARIB, CCSA, ETSI, ATIS, TTA and TTC and researches and development engineers from all the world

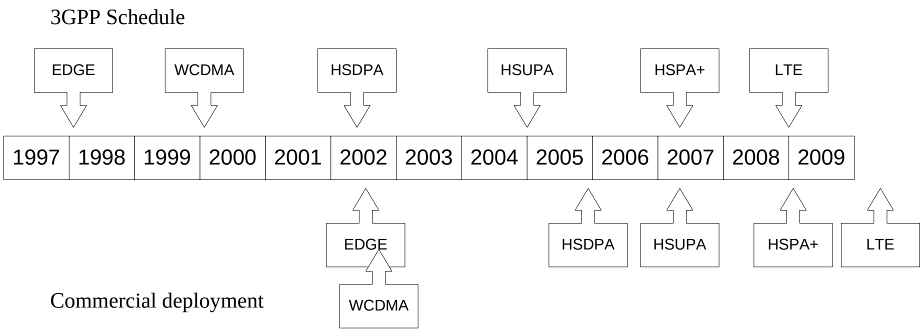


Figure 2.1: 3GPP schedule for standard releases and deployments

- Significantly improved spectrum efficiency (e.g. 2-4× Rel6 HSPA)
- Possibility for a Radio-access network latency (user-plane UE - RNC (or corresponding node above Node B) - UE) below 10 ms
- Significantly reduced control plane latency
- Scalable bandwidth (1.25, 2.5, 5, 10, 20 and possibly 15 MHz)
- Support for inter-working with existing 3G systems and non-3GPP specified systems
- Further enhanced MBMS
- Reduced CAPEX and OPEX including back haul
- Cost effective migration from Rel-6 UTRA radio interface and architecture
- Reasonable system and terminal complexity, cost, and power consumption
- Support of further enhanced IMS and core network
- Backwards compatibility, but with carefully considered trade off versus performance and/or capability enhancements
- Efficient support of the various types of services, especially from the PS domain (e.g. Voice over IP, Presence)
- System optimized for low mobile speed but also supported high mobile speed
- Operation in paired and unpaired spectrum not precluded
- Possibility for simplified co-existence between operators in adjacent bands as well as cross-border co-existence

The approved results satisfy and in some cases exceed some of the proposed targets in term of data rates, throughput and spectrum efficiency [27].

2.1 Background

A lot of specifications for LTE architectures and technologies have been studied to comply a large set of scenarios as depicted in [4]:

- *Standalone deployment scenarios* where operators deploy E-UTRAN in areas without a previous network or particular cases where coverage does not require internetworking with existing UTRAN/GSM EDGE Radio Access Network(GERAN)
- *Integration with existing UTRAN and/or GERAN deployment scenarios* where operators already have UTRAN and/or GERAN networks with any level of coverage and maturity

The objectives of E-UTRAN involve the easy deployment of new technologies in the previously described scenarios, supporting shared networks, high-velocity and nomadic mobiles with short latency and low data loss, different cell sizes and radio environments, support and cooperation with legacy systems and obviously radio efficiency.

LTE standardization procedures have defined specifications for both air interface attributes and network architecture (figure 2.2).

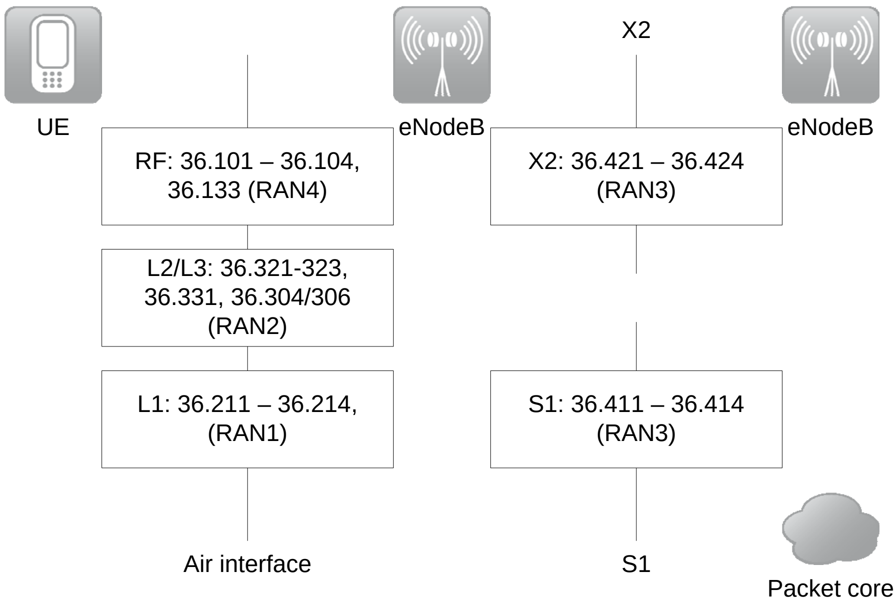


Figure 2.2: E-UTRA/E-UTRAN layers and relative specifications

The former are listed in table 2.1. OFDMA and SC-FDMA modulations allow flexible bandwidths, joint with Frequency Division Multiplexing (FDD) and Time Division Multiplexing (TDD). The service is optimized for low speed up to 15Km/h, however speeds up to 350Km/h are supported with performance degradations. The peak data rates are of 326Mb/s with 4 x 4 Multiple Input Multiple Output (MIMO) antennas configuration, using 20MHz bandwidth providing a cell spectral efficiency four times higher compared to Release 6 HSPA system. The transmission latency is also improved to smaller than 10ms values [4].

LTE meets International Mobile Telecommunications-2000 (IMT-2000) requirements and hence it is also part of IMT-2000 family of standards.

Table 2.1: LTE air interface specifications

Bandwidth		1.25-20MHz
Duplexing		FDD, TDD, half-duplex FDD [7]
Supported UE Speed		up to 350 km/h
Multiple access	Downlink	OFDMA
	Uplink	SC-FDMA
MIMO (UE class dependent)	Downlink	2×2, 4×2, 4×4
	Uplink	1×2, 1×4
Peak data rate in 20MHz	Downlink	172 and 326 Mb/s for 2×2 and 4×4 MIMO, respectively
	Uplink	86 Mb/s with 1×2 antenna configuration
Modulation		QPSK, 16-QAM and 64-QAM
Channel coding		Turbo code
Other optional techniques		Channel sensitive scheduling, link adaptation, power control, ICIC and hybrid ARQ

Figure 2.3 shows the proposed E-UTRAN architecture, with blocks of both control and user planes.

Jointly to new air interface architecture 3GPP is standardizing an evolution of the packet core network too (figure 2.4). The objective is to build an Internet Protocol (IP) based, flat network architecture with a low level of complexity. This new architecture is referred to as System Architecture Evolution (SAE), and represent an improvement of existing GSM/Wideband Code Division Multiple Access (WCDMA) core networks, designed to support mass-market use of IP-based services with more cost-efficient deployment and optimized network performance. The SAE architecture specifies only two types of node, the LTE base station E-UTRAN Node B (eNodeB) and the SAE Gateway: the connection between eNodeB and Core Network is realized with the S1 interface.

Also in the core network architecture, 3GPP efforts to build a network compatible with existing ones result in the connection of GSM, WCDMA, HSPA, CDMA2000 1xRTT, EV-DO systems with SAE by mean of specific and standardized interfaces. One of the major benefits is the possibility of both dual and single radio handovers, allowing a coexistence with and easy migration from existing systems. Other fulfilled objectives are:

Control signaling efficiency achieved making a separate system, the Mobility Management Entity (MME), to handle it, improving the capacity scaling abilities.

IP-based signaling and charging obtained abandoning the Signaling System 7 (SS7) of GSM and WCDMA in favor of Diameter.

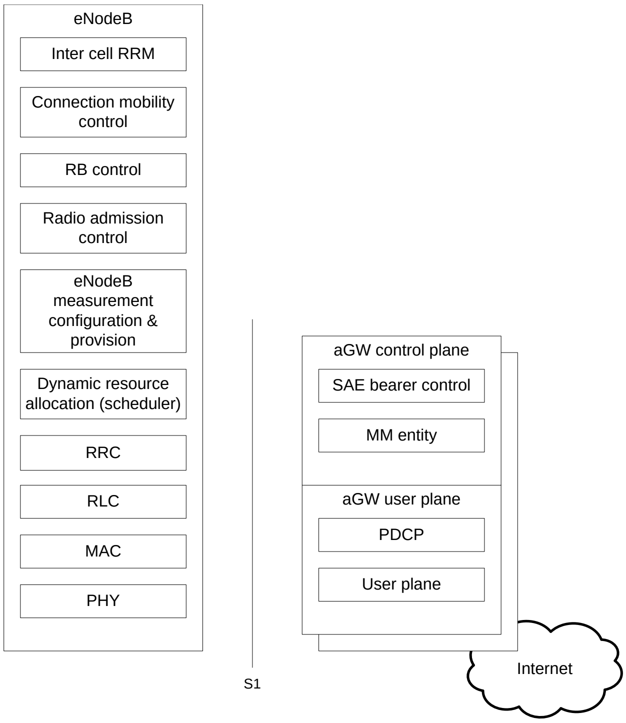


Figure 2.3: E-UTRAN architecture with control and user planes

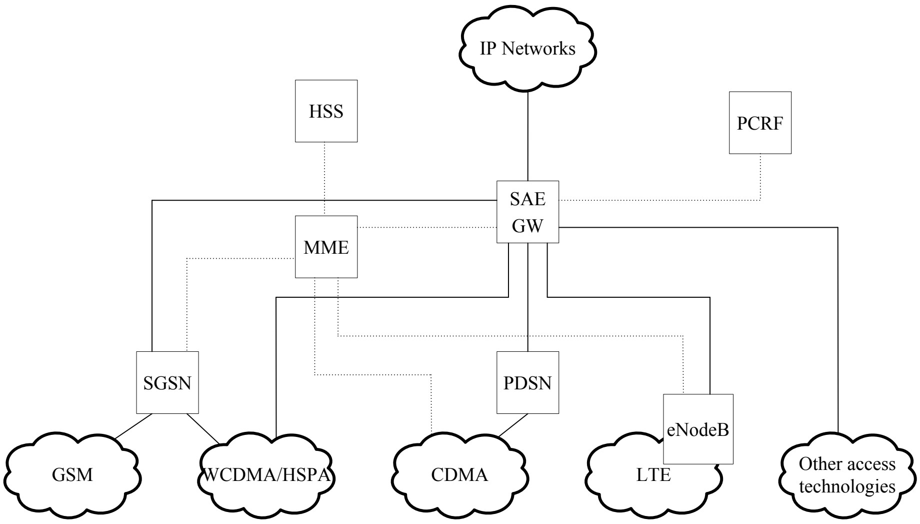


Figure 2.4: LTE and SAE combined architecture

Class-based QoS exploitable by operators to offer differentiated typologies of packet-based services, with Guarantee Bit Rate (GBR)

2.2 Radio Access

The radio access technology used in 3G system, known as WCDMA makes use of orthogonal Walsh codes to scramble radio channels and limit the inter-user and inter-symbol interference (ISI). However, this technique used on 5MHz bandwidths already suffers from multi-path interference phenomena, that worsen if scaled up to 10MHz or 20MHz [16]. Assumed this, it has become clear that LTE had to use a different Radio Access Technology (RAT) to achieve proposed requirements and offer great scalability attributes.

2.2.1 OFDMA

Transmission of data over a Single Carrier (SC) are performed modulating the input signal on an appropriate waveform (usually a sinusoid), varying the instantaneous value of amplitude, frequency or phase of it. The carrier frequency is usually much higher then the signal one. With Quadrature Amplitude Modulation (QAM) the transmitter adjusts the phase and amplitude of two waves, 90 degrees out-of-phase with each other (i.e. in quadrature) according to a specified symbol (dependent on QAM constellation order), that brings a certain number of bits. The value of data-rate impacts on the required bandwidth among the RF spectrum. Sub-carriers are already modulated carriers that act like input signal on new modulation processes.

The Frequency Division Multiple Access (FDMA) scheme allows the transmission of different user signals over different carriers or sub-carriers. Each transmission bandwidth center is located by a filter bank in a specified point of the spectrum making multiple users simultaneously and efficiently exploit the available spectrum. Obviously, Carrier distances on frequency domain have to be finely set to avoid interferences between signals. An efficient solution that avoids inter-signal interferences between adjacent sub-carriers and simultaneously allows a little overlap between them, is represented by Orthogonal Frequency Division Multiple Access (OFDMA) . This technique is based on the selection of adjacent sub-carriers whose centers are distant in a way that each of neighboring ones have zero value in frequency domain during the sampling instant of the selected one. Chosen sub-carriers are referred to as “orthogonal” and the resultant signal is strongly resistant to multipath losses. The principle of sub-carriers orthogonality is known since 1950s, but only modern digital equipments can keep this property between sub-carriers without being affected by components and temperature variations. Currently OFDMA is successfully used in many areas of digital transmissions, like Digital Video Broadcasting - Terrestrial (DVB-T) , Digital Video Broadcasting - Handheld (DVB-H) and WLAN, but is relatively new in cellular systems.

LTE standardization chose OFDMA because of its benefits, that include high performances with low cost and complexity devices, efficient spectrum usage and compatibility with advanced antenna and receiver technologies.

The implementation of OFDMA in E-UTRAN is based on Discrete Fourier Transform (DFT) and the inverse operation (IDFT) that can switch the representation of a signal between time and frequency domains. The Fast Fourier Transform (FFT) implementation sized in power of two blocks, joint to classic requirements for signal processing in terms of sampling frequency and words length, can assure a lossless transmission of the original signal. Furthermore LTE specifications fix the sub-carriers spacing to 15kHz to make the system tolerate frequency offsets due to Doppler shift produced by UE speed and electrical components imperfections, and time domain guard intervals to counter the delay spread of the deployment environment. OFDMA transmission, however, suffers from inter-cell frequency interference when user is located at the cell edge. CDMA system are less prone to this because different scrambling codes are used in adjacent cells. An efficient plan of frequency reuse among near cells can limit this issue, but LTE Intercell Interference Coordination can make the system to work efficiently with a frequency reuse factor of 1 too.

The OFDMA transmissions therefore result in several parallel sub-carriers in time domain, a set of sinusoidal waves that fill the bandwidth with 15kHz steps in time domain. This suggests a representation of the frequency/time domains as a grid, in which LTE specifies so-called Resource Blocks (RB) made of 12 sub-carriers of 15kHz that last for 1ms (figure 2.5).

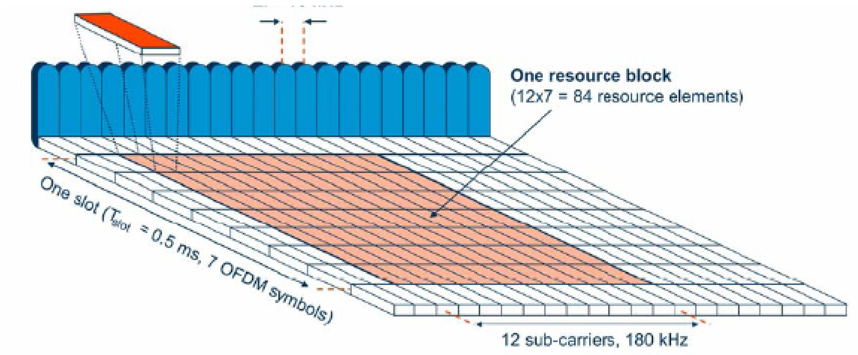


Figure 2.5: T/F Grid with Resource Block (RB) representation

The number of RB assigned to users depends on modulation and coding scheme supported, and the correct allocation of them among the grid is made by time and frequency domains scheduling algorithms.

One of the side effect of allowing an high number of parallel transmission is caused by high and fast variability of the signal in time domain, with an excessive Peak To Average Ratio (PAR) value. This implies some drawbacks in the design of power amplifiers. This is why LTE uplink radio access is realized with a different technique.

2.2.2 SC-FDMA

OFDMA transmissions implies high PAR values in signal, that can bring the amplifiers to very low efficiency values, and high power consumption. Uplink radio access for LTE is based on Single Carrier FDMA (SC-FDMA) to avoid this issue. SC-FDMA combine low PAR techniques of single-carrier transmissions, such GSM and CDMA with the multi-path resistance and flexible frequency allocation of OFDMA [8]. In SC-FDMA data symbols in time domain are converted to the frequency one using DFT, then they are mapped to the desired location of the available bandwidth and finally reconverted with IFFT in time domain adding in the meanwhile the cycling prefix (CP). SC-FDMA is also referred to as DFT spread OFDM (DFT-SOFDM), because of this processes. The practical difference between OFDMA and SC-FDMA consists in the fact that the former transmits modulation symbols in parallel over multiple sub-carriers, beside SC-FDMA transmit a symbol over all sub-carriers more times the rate of OFDMA. Each data symbol seems to be transmitted by one single carrier, and this is the SC reason of the technique name. The serial transmission of symbols contributes to lower the PAR value to the original data symbols one, occupying the same bandwidth of multi-carrier OFDMA. The generation of SC-FDMA symbols involves a special pre-coding of data symbols, and it can be demonstrated that it makes this technique resistant to multipath as well as OFDMA [8].

2.2.3 MIMO Spatial Multiplexing

Transmission diversity obtained with multiple transmission and reception antennas can be used to achieved high diversity gain. The transmission diversity can be exploited to improve system performances, especially when UE's conditions make channel quality feedback reporting unreliable, or impossible, i.e. when UE's speed is high. It's possible to improve peak data rates enabling multiple data streams between multiple transmission antennas at the eNodeB and reception antennas at the UE. LTE supports three multiple antenna schemes: Transmission diversity with Multiple Input Single Output (MISO), Reception diversity with Single Input Multiple Output (SIMO) and spatial multiplexing with Multiple Input Multiple Output (MIMO).

2.3 Physical channel and frame structures

One of the main feature planned when LTE has been proposed was the support for flexible bandwidth, up to 20MHz. The physical layer architecture therefore has been built to support bandwidth increments of 180kHz starting from a bandwidth of 1.25MHz [16]. Furthermore, to achieve low latency transmissions, Transmission Time Interval (TTI) has been fixed to 1ms. As stated in 2.2.1 sub-carriers are spaced of 15kHz, to ensure robustness to Doppler spread effect, but a smaller value of 7.5kHz is also standardized to support large delay spreads. Uplink transmission are performed in SC mode, so they have to be allocated on adjacent sets of sub-carriers.

LTE supports six channel bandwidths (Table 2.2), with different numbers of resource block allocable. It's mandatory to observe that the transmission bandwidth is smaller than channel bandwidth, in fact BW_{config} is 90% of the channel bandwidth $BW_{channel}$. The relationship between BW_{config} and transmission configuration N_{RB} is given by equation (2.1).

$$BW_{config} = \left(\frac{N_{RB} \times N_{SC}^{RB} \times \Delta f}{1000} \right) \tag{2.1}$$

where sub-carriers spacing $\Delta f = 15kHz$

Table 2.2: Transmission bandwidth and configuration for LTE six channel bandwidths

Channel bandwidth $BW_{channel}$ [MHz]	Transmission configuration N_{RB} [#RB]	Transmission bandwidth BW_{config} [MHz]
1.4	6	1.08
3	15	2.7
5	25	4.5
10	50	9
15	75	13.5
20	100	18

2.3.1 Frame and slot structure

Both downlink and uplink LTE transmission are based on a 1ms scheduling granularity. The fundamental subframe lasts for 1 ms and includes two 0.5ms time slots. Each of system control signal is performed with a specified subframe periodicity, for example synchronization and broadcast signals are transmitted on a radio frame basis i.e. 10 sub-frames. Each time slot is divided in N_{symb}^{DL} symbols for OFDMA downlink and N_{symb}^{UL} for SC-FDMA uplink. Figure 2.6 summarizes the previous assertions.

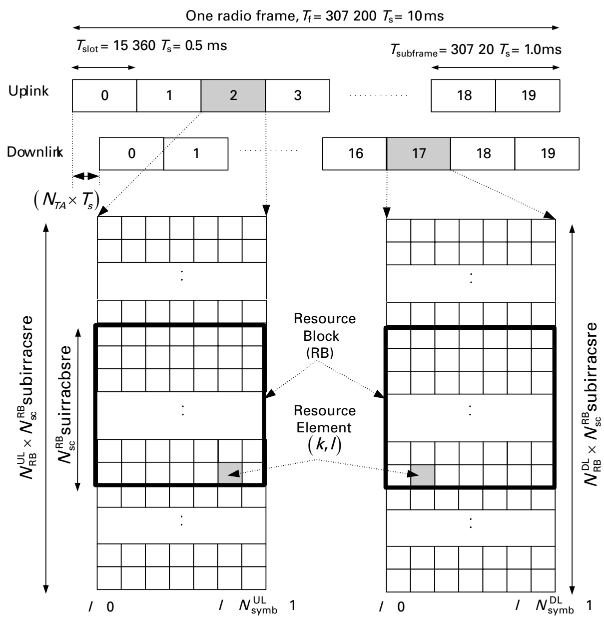


Figure 2.6: Frame structure [13]

A Physical Resource Block (PRB) is a set of N_{SC}^{RB} adjacent sub-carriers in frequency domain and N_{symb}^{DL} or N_{symb}^{UL} symbols for downlink and uplink respectively. LTE specifies three configurations for these parameters [2] (Table 2.3).

Table 2.3: Physical Resource Block configurations

Configuration		N_{SC}^{RB}	N_{symb}^{DL}	N_{symb}^{UL}
Normal cyclic prefix	$\Delta f = 15kHz$	12	7	7
Extended cyclic prefix	$\Delta f = 15kHz$	12	6	6
Extended cyclic prefix	$\Delta f = 7.5kHz$	24	3	NA

2.4 RRM

Radio resource management is the set of system level controls principally run by eNodeB, that ensure an efficient utilization of the radio channel, in order to ensure both stability, a working system and optimise user and system throughputs. In the design of a cellular wireless communication system there is always a trade-off between UE complexity, network complexity and achievable performances [27].

LTE RRM framework includes a wide series of algorithms that provide different control services enclosed in main categories like transmission power control, allocation of resources and QoS management. All these services work among several layers of ISO/OSI stack, from Layer 1 with power control and channel estimation, to Layer 3 with QoS functionalities (figure2.7), but they are strictly interconnected to ensure the best use of the available resources.

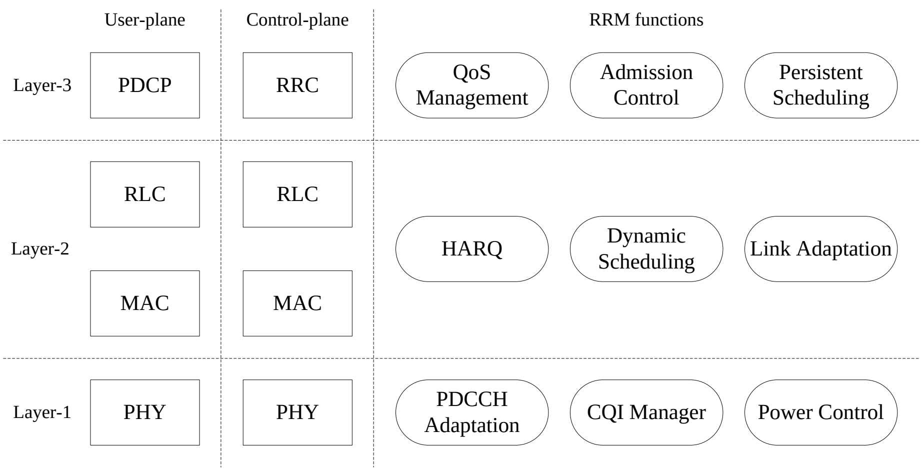


Figure 2.7: Overview of eNodeB user and control plane, with functionalities mapping over ISO/OSI layers

[2]Sub-carriers spacing of 7.5kHz is only used for Multicast Broadcast Single Frequency Network (MBSFN)

RRC protocols involve transmissions of common Non-Access Stratum (NAS) informations (equals for all UEs) and dedicated ones, and they can be divided in generic areas [27]:

- *System information* concerns the broadcasting of common informations for UEs either in RRC_CONNECTED state or RRC_IDLE
- *RRC connection management* involves all the procedures related to RRC connections control, paging, security features, handover, access class barring, radio link failures and lower layers parameters setting
- *Network controlled inter-RAT mobility* covers security activation and is responsible for UEs RRC context information transfer
- *measurement informations and reporting* for inter-RAT, intra-frequency and inter-frequency mobility that involves configuration and activation of measurement entities
- *miscellaneous functions* especially for delivering of dedicated NAS informations

Figure illustrates the whole LTE radio architecture with the mapping of several layers channels and functionalities.

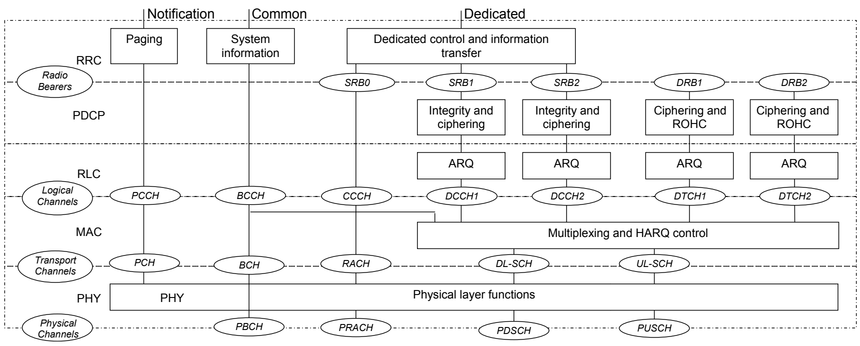


Figure 2.8: Radio Architecture for LTE. All channels denominations are defined in [6], [5] and [7]

3GPP specifies two radio resource control states for UE [7], and related specific protocols to handle UEs in each state:

RRC_IDLE UEs in RRC_IDLE state show sporadic activity that mainly involves cell selection and reselection: UE chooses the cell basing its decision on priority of each applicable frequency of each RAT, radio channel quality cell status (avoids barred or reserved cells). In RRC_IDLE the UE listens to paging channel too, receiving informations about incoming calls, and system parameters specified by E-UTRAN that assist in the cell selection process.

RRC_CONNECTED UEs in RRC_CONNECTED state have allocated radio resources in shared data channels, specified by dedicated signaling performed via control channels. Also, in this state, the UE periodically reports downlink channel quality to eNB, as well as neighboring cells informations including cells using other frequencies or RATs.

In the following sections we focus the available choices about one of the fundamental features of RRM that involves UEs in RRC_CONNECTED state, the packet scheduling. Indeed, 3GPP specifies the RRM related signaling, leaving the scheduling algorithms choices to vendors and operators [13].

2.4.1 Signaling

Every transmission performed in DL and UL data channel, PDSCH (Physical Downlink Shared Channel) and PUSCH (Physical Uplink Shared Channel) respectively, requires a certain amount of signaling on control channels PDCCH (Physical Downlink Control Channel) and PUCCH (Physical Uplink Control Channel). The signaling includes control information originated by physical (L1) and MAC (L2) layers. The DL control informations involve scheduling information and MCS choices required to correctly receive, demodulate and decode PDSCH data for each scheduled terminal, as well as information on some transmission over PUSCH, i.e. scheduling grants. The UL signaling includes HARQ acknowledgments for DL received packets, channel quality indicators to fully exploit time and frequency diversity gains and Scheduling Requests needed in case of available UL data.

The biggest load of control channel is due to CQI reporting, and especially with VoIP traffic it has to be correctly configured to avoid to push the system into a situation where control channels are fully used and a lot of data channel resources are unused because no more UE can be controlled. Indeed, control channels overhead can be one of the main limiting factor for VoIP capacity.

2.4.2 Packet scheduler structure

LTE networks are completely "packet oriented" with a packet switched core network, so packet scheduling has a key role to achieve high QoS levels with high spectral efficiency in this multi-user environment. Optimizing a wireless network system for VoIP traffic, taking into account channel conditions while making scheduling decisions, can boost the whole capacity up to 40% [26].

Designing a packet scheduler for VoIP traffic should consider some special purposes due to its nature: *(a)* delay constraints of VoIP service — as said in chapter 1 the air interface delay should be kept to values lower than 150ms to avoid the user's call experience to degrade rapidly. LTE specifications try to go beyond this threshold, fixing a maximum value of 50ms [13]. This implies that the scheduler must consider the eventuality of retransmissions: users that experience poor channel quality will cause a number of HARQ retransmissions that can possibly break the maximum delay threshold, thus the scheduler has to take into account delay budget and retransmission togheter. *(b)* low data rate and constant periodicity — Every VoIP packet has a fixed size, typically 320 bits with Robust Header Compression (ROHC), produced every 20ms for AMR 12.2 codec. These characteristics imply that the small payloads have to be scheduled very often, with the possibility to group them only in case of good channel conditions [13]. *(c)* administration of control channel usage — with a fully dynamic scheduling policy, the main bottleneck of VoIP service is the heavy usage of control channels: in heavy load conditions, it's possible to experience lots of unused PDSCH resources with a completely exhausted PDCCH [13]. This can be avoided with channel quality feedback based techniques like packet bundling, but only if the UE is experiencing favourable channel conditions (i.e. UE not at cell edge)[13].

PS regularly performs decisions about allocation of PRBs to users that require resources in a certain time instant. Furthermore it has strong connections with link adaptation and HARQ management entities, to ensure per user based QoS requirements. Main relations of the packet scheduler with other RRC entities are [13]:

packet scheduler with DRX/DTX controller select only users that are not in power saving state

packet scheduler with power control the RF modem activation must be managed according to scheduled transmission instants

packet scheduler and QoS packet scheduler is one of key factors to respect QoS constraints

packet scheduler and Buffer Status Report (BSR) packet scheduler has to take into account data buffer statuses, and priority relations among multiple users buffers

packet scheduler and HARQ packet scheduler has to take into account the possibility of HARQ retransmissions in the utilization of delay budget, to fulfill QoS requirements

Unlike CDMA systems, LTE is a TDMA/FDMA system: the elementary block allocable to a user (Resource Unit RU) consists of 12 OFDMA adjacent sub-carriers of 15kHz for 1ms. The purpose of the scheduler is to maximize the efficiency of utilization of these RU-TTI. Due to the combined utilization of TDMA and FDMA, the LTE PS can be divided in two main, independently configurable blocks that perform subsequent choices in time and frequency domains. Thanks to this, opposed to joint time/domain schedulers that achieve almost the same performance with higher complexity, LTE scheduler runs faster due to the reduced number of users the frequency domain scheduler has to manage [16]. The performances of decoupling time/domain scheduling for QoS management have been proved in [21].

The scheduling algorithms take into account the Channel Quality Indicators (CQI) reporting, as well as the link adaptation dynamics and the informations reported by the HARQ manager.

Time domain scheduler

First part of LTE scheduling process runs every TTI and involves the selection of N users out of the total amount scheduled for the next TTI, called Scheduling Candidate Set (SCS). This choice is based on several conditions and metrics: among all UE with pending data, the time domain scheduler has to find which one is not under DRX/DTX conditions, and apply an algorithm to find out

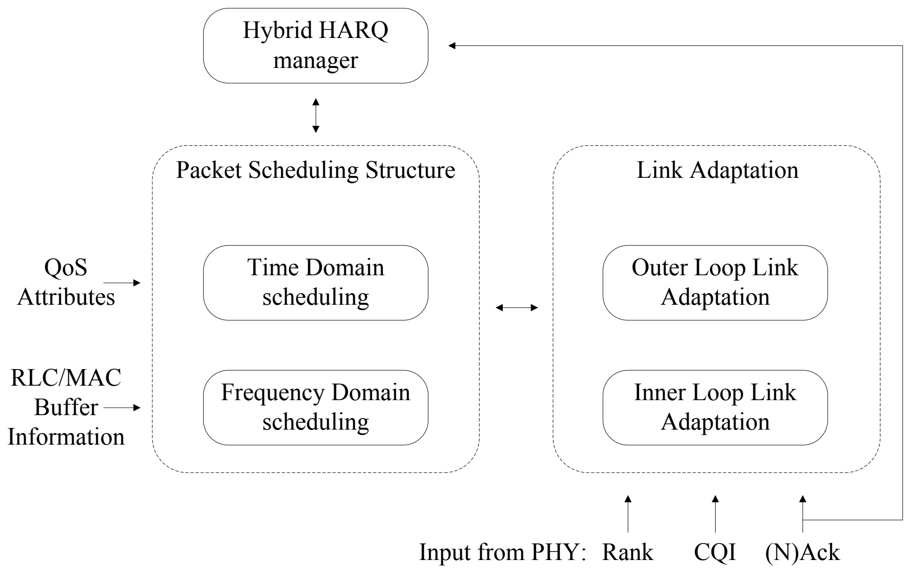


Figure 2.9: Time/frequency schedulers

the candidate users to achieve QoS constraints based on prioritization metrics. Furthermore, it has to estimate the total amount of control channels capacity to be able to control the maximum number of UEs. The proportional fair algorithm extended with the OFDMA time/domain specialized nature of LTE, exhibits high throughput [14] [28] and [18] works as a part of the admission control system.

The time domain scheduler takes into account the pending HARQ retransmissions as well as normal transmissions. It gives an higher precedence to the former ones, since it's not possible to combine them in a single TTI (for one user, of course) [5]. However, under particular circumstances, it's possible to bundle more data.

Frequency domain scheduler

Once the total amount of users scheduled for successive TTI is reduced to a set of candidates by time domain scheduling, the frequency domain scheduler distributes the available resources units among them guaranteeing a certain level of fairness among transmission and HARQ retransmissions, and above all, trying to exploit the channel conditions instantaneously experienced by each user, to be able to assign PRBs located on those frequencies where he perceives the higher channel quality. In fact, the UE experiences wide power variations in both signal and interferences, due to radio signal power alteration phenomena like fading.

To be able to provide high gains in frequency diversity exploiting, the frequency domain packet scheduler needs an accurate estimation of the whole channel quality. Based on scheduling algorithm chosen, the Channel Quality Indicator can be reported in several ways.

It is important to underline that the uplink SC-FDMA technique puts some additional constraints compared to downlink in the scheduling choice of PRBs.

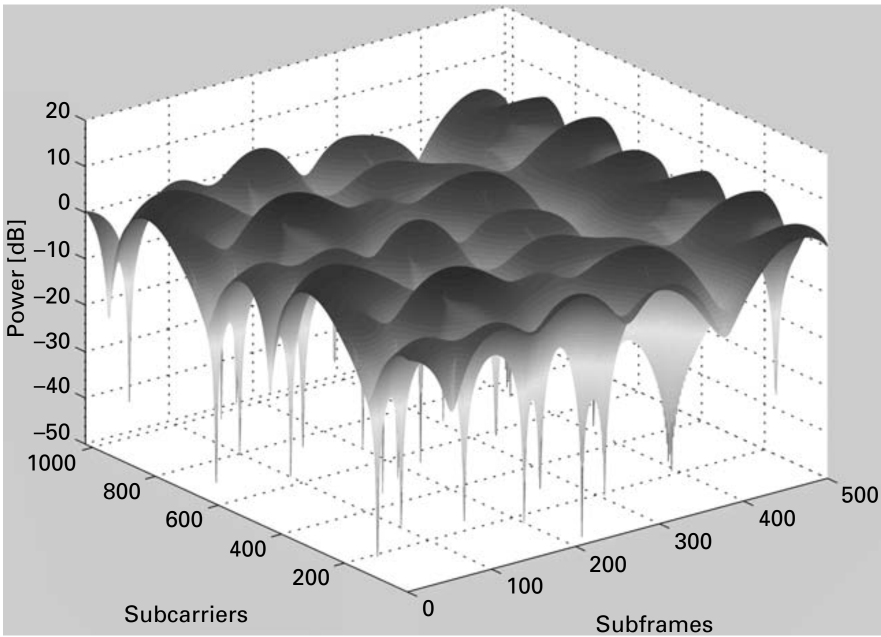


Figure 2.10: Signal power variation in time and frequency

In fact, the single carrier characterization forces the scheduler module to allocated continuous segments of PRBs among the whole bandwidth. The final result is a lot less fragmented allocation table among the scheduled users (figure 2.11).

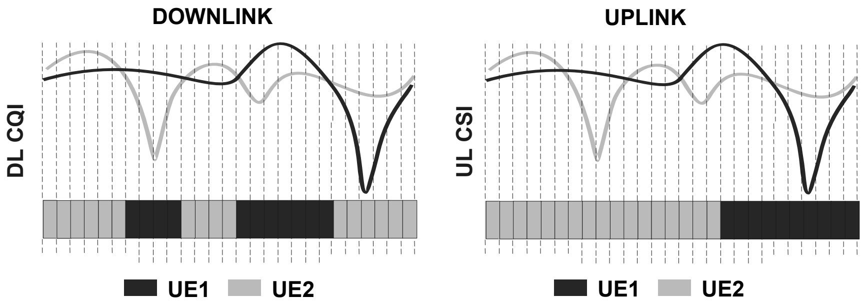


Figure 2.11: Frequency allocation comparison in DL and UL due to SC-FDMA constraints [13]

2.4.3 Packet scheduling policies

Several packet scheduling algorithms have been proposed in LTE, and especially to correct handle VoIP traffic. These have to find the correctly balance between user satisfaction (chapter 1) and efficient use of available data and control channels.

Dynamic packet scheduling

The packet based nature of LTE suggests that each transmission between eNodeB and UE can be dynamically scheduled each time to exploit both time and frequency diversity. Obviously this could be applied to VoIP traffic too, but it's mandatory to take into account that, although it allows flexibility and accuracy, it can require a huge amount of L1 and L2 signaling, i.e. a resource request for each VoIP packet. In fact a decision about transmission parameters, as well as modulation and coding scheme, is made by Layer 2 each TTI. To achieve high accuracy levels, it is possible to require a really frequent channel feedback report on control channel, it will be described in section 2.4.4, but, in this case the control overhead may become a limiting factor. It's important to specify that the scheduling decision is performed on a per user basis, even if the UE has several data flows [13].

The greatest benefit of this solution, it's the correct estimation of needed resources in terms of modulation and coding scheme along the whole duration of a talk spurt. In fact, if the frequency of the CQI reporting is appropriately chosen according to AMR codec frequency, it's possible to make a precise decision that exploits fresh and accurate values of channel quality feedback, limiting the number of needed retransmissions.

Semi-persistent packet scheduling

To avoid control channel capacity limitations, 3GPP has adopted the semi-persistent scheduling solution for voice type traffic [5]. The semi-persistent packet scheduling includes some elements of persistent one like assigning a fixed set of recurrent resource blocks that last for the whole talk activity, plus dynamic scheduling for retransmissions and SID. When the scheduler detects a new talk-spurt, it performs the choice about transport format, and regularly assigns resources with a period conveniently chosen in accordance with AMR periodicity [7]. The SPS decisions are performed by Layer 3 of the ISO/OSI stack, so they require a slight higher delay to be completed. The resources persistently allocated to the UE are implicitly released at the end of talk activity after a configurable amount of empty transmissions [7]. Of course it's possible to renegotiate the allocation to perform a rough link adaptation in presence of an high number of consecutive retransmissions.

The main downside of SPS are the less favorable performances given by the decoupling between a fixed MCS and highly variable channel conditions. With talk spurts length of 2s (table 3.4), it is easy to understand that the choice made at the beginning of a talk-spurt is based on channel conditions that last for a very short fraction of the talk activity. A possible method to reduce the impact of fast power variation of signal and interference during the talk-spurt length is to apply a preprocessing of the current CQI memory, to make more conservative choices. Underestimating the channel quality, thus lowering the modulation and increasing the coding level, makes possible to keep a low BLEP during the whole talk-spurt.

2.4.4 Channel quality feedback

In order to improve spectral efficiency and low transmission errors, scheduling is performed adapting the modulation and coding scheme of the transmission to the channel state between the eNodeB and UE. In downlink such information is not directly accessible by the eNodeB, so the UE must deliver a channel quality indicator (CQI) that consist of the highest MCS index that it can decode with a block error rate below say, 10%. This index is the result of measurements based on signal to interference-plus-noise ratio (SINR) estimated by listening to some reference symbols. To achieve maximum frequency diversity gains in DL, each UE should estimate the channel quality per PRB, across the whole bandwidth. However a signaling policy like that, especially with VoIP traffic where an high number of UE can be scheduled together, can lead to serious issues with the control channel capacity due the amount of UL control channels overhead that can reach 8% for 10MHz [13]. Hence, it is mandatory to reduce the channel quality feedback in time or frequency domains. The time domain reduction depends on the speed of the UE: with 3Km/h a periodicity of 5/10 ms should be a reasonable choice [9].

A simple method to reduce the control overhead in frequency domain is to set a rougher granularity of the estimation. The frequency resolution of the measurements is controlled by the eNodeB defining the number of contiguous physical resource blocks (PRBs) [27] to estimate. The main frequency selective schemes are:

- *Wideband Feedback*: only a single CQI value is sent for the whole system bandwidth [13]. This modality allows a reduction of the number of bits to report the channel condition, and therefore the signaling overhead, but at the same time reduces the user frequency selectiveness to zero.
- *Best-M Average*: the UE estimates the channel quality of small groups of PRBs and reports only a single value corresponding to an average of the M best ones. In this case the overhead depends on the number of selected PRBs.
- *Full CQI reporting*: the UE reports the channel quality of small groups of PRBs [6] increasing the frequency selectiveness accuracy but at the same time the signaling overhead.

It is possible to note that an important aspect of CQI reporting consists of the uplink resource usage and optimization in terms of signaling overhead. For this reason the eNodeB can select the periodicity of each UE CQI report, according with the network condition, choosing between periodic or aperiodic report [13]:

- *Periodic CQI reporting*: the eNodeB requires the CQI feedback on the PUCCH in pre-defined time instant. The period can be set to 2, 5, 10, 16, 20, 32, 40, 64, 80, 160 ms or Off [27] and is configured via higher layer signaling. The size of a single report allows to contain only little information about the frequency channel state [13].
- *Aperiodic CQI reporting*: the eNodeB can instruct the UE to send an individual CQI report [13] on the PUSCH when it needs precise channel state information.

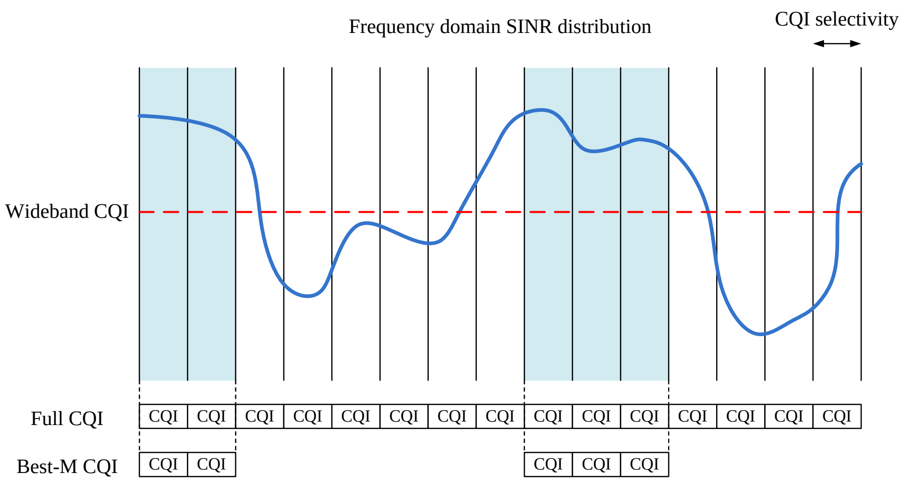


Figure 2.12: Illustration of the three CQI reporting methods

In uplink the channel quality is estimated by the eNodeB requiring sounding reference signals (SRS) from the UE. SRS can be considered a counterpart for the CQI reporting but does not contain any type of feedback information like the CQI. SRS is transmitted on the last symbol of the subframe on PUSCH as a single or periodic transmission, with period from 2 to 320 ms [13]. To provide flexibility, it is possible to choose different bandwidth options according with the UE power characteristics. The configuration parameters are sent via higher layer signaling.

2.4.5 LA and MCS

In general, in any cellular communication systems, the quality of the received signal is compromised by different phenomena inherent to wireless environments, like path loss, interferences, multipath propagation phenomenon. The objective of Link Adaptation is to adapt the resource allocation to the user channel condition, matching the transmission parameters such as modulation and coding scheme (MCS), pre-coding as well as transmission power control for physical channels [16], in order to avoid strong variations in the quality of received signal, to guarantee the required quality of service of each UE.

In the downlink transmissions, the eNodeB does not know the actual channel conditions of the UE, and for this reason, it can require a Channel Quality Indicator (CQI) feedback from the receiver to be assisted in selecting an appropriate MCS. The modulation scheme chosen for LTE consists in different-order QAM. In general, the eNodeB can select QPSK, 16-QAM and 64-QAM schemes and different code rates [27] to achieve a trade-off between high data throughput and low block error rate (BLER). Choosing a low-order modulation, the eNodeB guarantees a more robust transmission but a lower bit rate. Contrary, choosing high-order modulation the eNodeB allows higher data rate but lower robustness. The code rate adapts the chosen modulation scheme to the channel conditions in order to increase transmission reliability.

For uplink transmissions, the eNodeB handles the link adaptation in the same way of the downlink case, but instead of requiring CQI feedback, it can self-estimate the supportable uplink data rate using Sounding Reference Signals (SRSs) or base it on the past measured SINR. QPSK and 16-QAM can be selected as modulation schemes and, for highest category of UE, also 64-QAM [27].

2.4.6 Outer Loop Link Adaptation

The purpose of Outer Loop Link Adaptation (OLLA) is to make accurate scheduling choices based on CQI reported values. In fact, scheduling policies based only on these values, without any type of feedback control, are affected by an amplified impact of quantization and reporting errors, thus don't follow the expected BLEP target. This is mainly due to the variations of the SINR between the instant when the CQI is estimated, and the one when it's used to perform the scheduling decision.

Therefore the eNB applies a link adaptation algorithm that processes the received CQI index before using it [25]. This quality control of transmissions is based on the dynamics of ACK/NACK receptions of past ones. The schematic functionality is given by an adjustable offset A applied to the perceived CQI measure. If the first transmission receives an ACK, then the offset is decreased by A_{down} dB (equation (2.2)).

$$A = A - A_{down} \tag{2.2}$$

If a NACK is received, it's increased by A_{up} dB (equation (2.3)).

$$A = A + A_{up} \tag{2.3}$$

The CQI reports consist in an index word, so the eNB has to know the real dB differences between CQI indexes. This is realized with a fixed conversion table defined by 3GPP; each entry is denoted by $f_{SINR}(n)$ where n is the received CQI index and $\hat{n}$ the processed one [3]. The CQI index processed by link adaptation is:

$$\hat{n} = \arg\min_{\hat{n}} \{f_{SINR}(\hat{n}) - (f_{SINR}(n) - A)\} \tag{2.4}$$

With this approach it can be seen that the BLEP o first transmission is

$$BLEP = \frac{1}{1 + \frac{A_{up}}{A_{down}}} \tag{2.5}$$

Controlling the ratio between A_{up} and A_{down}, results in controlling the transmission BLEP, actuating more conservative choices in presence of high number of NACKs, i.e. an high number of retransmissions.

2.4.7 HARQ

LTE adopts a physical layer retransmission combining, called Hybrid Automatic Repeat Request (HARQ), to quickly and reliably provide data at physical layer,

[3]The tables defined by 3GPP follow the approximate property: $f_{SINR}(n) \simeq 1 \cdot n$ dB $\forall n \in [0 - 31]$, i.e., the resolution of the CQI mapping table is approximately 1 dB.

optimizing, at the same time, the network resources use. The HARQ uses stop-and-wait (SAW) protocol. After a transmission, the transmitting entity stops and waits for a positive or negative acknowledgment (ACK/NACK) before transmitting a new packet or retransmitting the same one.

The HARQ entity must handle transmission, reception and retransmission of transport blocks, and generation, reception and processing of ACK/NACK signaling. During reception activity, the HARQ receiver entity tries to decode the received packet and sends a feedback in the PUCCH. For negative acknowledgment (NACK), the erroneous packet is saved on the receiver HARQ soft buffer [5] waiting for a new copy to be send by the HARQ transmitter entity. The two copies of the same packet are combined using chase combining (CC) or incremental redundancy (IR) to obtain a single packet to be processed by the decoder. With CC the retransmitted packet is exactly the same packet of the first transmission [16]. By contrast, with incremental redundancy the retransmission has different rate matching parameters [16].

Eight independent stop-and-wait HARQ parallel process can be active at the same time in order to support continuous transmission. Each process handles a separate stop-and-wait operation and an own buffer.

LTE supports two HARQ schemes: asynchronous adaptive for the downlink retransmission and synchronous adaptive or non-adaptive for uplink retransmission. In synchronous scheme the retransmission can occur only in predefined instants of time. So the eNodeB knows exactly when and which HARQ process must be processed. In this way the scheduling signaling is not necessary if it is on the same PRBs. In contrast, in asynchronous scheme the retransmission can be scheduled in any instant to the initial transmission and for this reason an HARQ process identifier must be sent to associate the retransmission with the original transmission. Adaptive or non-adaptive means that in retransmission MCS and resource allocation can or can not be changed at each retransmission [27].

2.5 DRX/DTX functionalities

LTE provides a set of functionalities to make UEs to perform micro sleep events either in RRC_IDLE or RRC_CONNECTED state, to be able to extend battery life though guaranteeing high QoS and connectivity. The actual implementation allows the UE to not constantly monitor the PDCCH every TTI, but only during specific time interval set by higher layers. This solution provides benefits both in downlink and uplink because all the scheduling control informations are transmitted on PDCCH. During non active states, the UE can go into power saving states that dramatically decrease the power consumption impact of the RF modem. Indeed LTE should be the reply to the oncoming desire of continuous connectivity, where the user remains persistently connected to the network for long time spans, with a generally low activity frequency of transmission and reception, avoiding numerous connection establishments and releases that induce overhead and delay [22].

The RRC plays a crucial role in the DRX/DTX management since it performs the biggest part of parameters setting for each UE. These parameters are listed in table 2.4.

The DRX functionality in RRC_CONNECTED state provides two DRX

Table 2.4: DRX parameters

DRX Parameter	Description
DRX Cycle	Specifies the periodic repetition of an active state that last for on Duration, followed by an inactive period (figure 2.13)
On Duration timer	Specifies how many subframes the UE should be in active state when a new DRX cycle starts (figure 2.13)
DRX Inactivity timer	Specifies how many PDCCH subframes after successfully decoding a PDCCH the UE must remain active. It's a simple method to keep the UE alive, especially useful for bursty traffic, not for regular one (like VoIP) (figure 2.13)
DRX Retransmission timer	Specifies the maximum number of consecutive PDCCH subframes the UE should remain active to wait an incoming retransmission after the first available retransmission time. Very useful for asynchronous HARQ (figure 2.15)
DRX Short Cycle	Specifies the periodic repetition of an active state when the UE is under short DRX condition, it's a sort of discontinuous inactivity timer (figure 2.14)
DRX Short Cycle timer	Specifies the consecutive number of subframes the UE must the UE shall follow the short DRX cycle after the DRX Inactivity Timer has expired (figure 2.14)

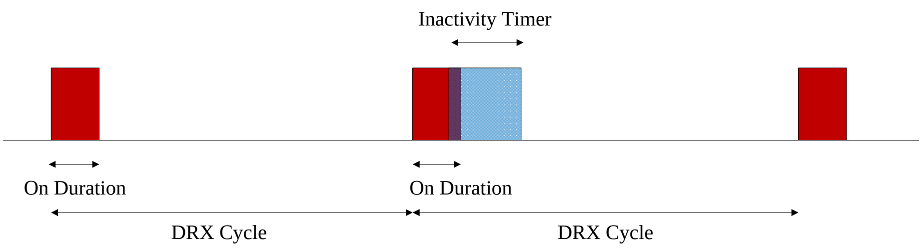


Figure 2.13: DRX states

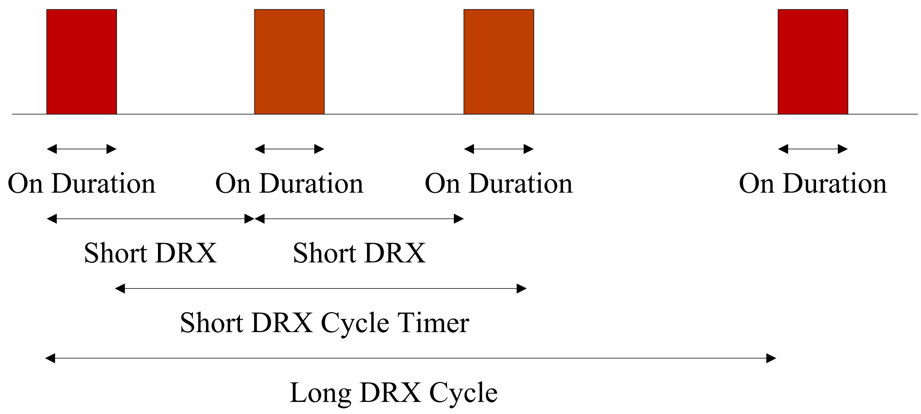


Figure 2.14: Long DRX Cycle and Short DRX Cycle

cycles that can be set for each UE. The long DRX cycle is used during UE's inactivity periods, when it only has to check the control channels and no resources are assigned: every *DRX Cycle* the RF modem is turned on for *On Duration timer* consecutive subframes to listen to the control channels. When data activity is detected both in downlink or uplink, the eNB triggers the short DRX cycle for UE, increasing the responsiveness and connectivity of UE. The transitions between long and short drx schemes are trigged directly by eNB, or timer driven.

Another DRX feature is related to the power saving during HARQ retransmissions: when the UE fails to decode a transport block of one of HARQ active processes, it assumes that the next retransmission will take place at least after *DRX Retransmission timer*, so it can go in power saving state without the needs to listen to PDCCH.

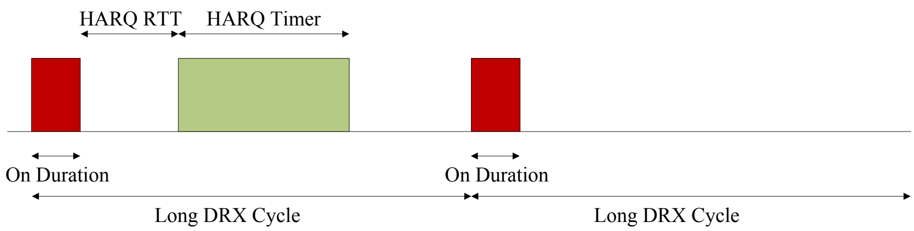


Figure 2.15: DRX states and HARQ retransmissions

It's notable to understand that the RRC should take its DRX/DTX decisions finding the best trade-off between power saving and responsiveness of the UEs. There are numerous DRX/DTX policies implementable to find the correct balance between those two purposes, and each one shows a different impact in term of:

- flexibility and complexity of the solution
- entities that can control DRX/DTX parameters at various levels, taking into account that decisions made by higher layers imply higher delays and logic executions that often don't worth doing them.

- periodicity of parameters changes to allow a fast adaptation of the DRX activity to the traffic parameters

Obviously, having an *a priori* knowledge of the traffic transferred between eNodeB and UE can represent the key to reach high performance with a parameters set completely fitted to the data dynamics. For example, with an high bursty traffic, like web browsing, it's useless to constantly monitor the shared data channel while user is reading a page, and can be useful to trigger a short DRX period when user fetch another page, to provide a faster response. VoIP traffic is high predictable and it shows a dynamic that can be fully exploited when setting DRX parameters. For example, the On Duration timer can be set to 1 subframe with a short DRX cycle of 20ms to totally fit the AMR codec output. HARQ retransmission are not taken into account with DRX cycles because they are performed at random instant that would force the DRX parameters setting to worst case scenarios.

Also the CQI and SRS reporting functionality is bound to DRX parameters, and takes power saving advantages when its periodicity and phase is equal to DRX one.

For LTE, DRX could represent an important success factor due to its power saving benefits, but it has to be tightly connected with all other RRC procedures, since it actually takes out an user from the scheduling candidates set thus it can reduce the multi-user packet scheduling gains. So it's mandatory to set correct DRX parameters that consider the multi-user scenario and user related QoS needs, like latency and responsiveness.

Chapter 3

Modeling

In this chapter we will analyze the actual implementation of the LTE features described in chapter 2, and how we have modeled some of the system entities.

3.1 Simulator architecture

As previously introduced, the developed simulator is a MATLAB®program, that performs system level analysis working on two link level traces provided by two Nokia Siemens Networks link level simulators. The script can handle a single user simulation, with a fine tuned parametrization of the main characteristics involved in our analysis. The main structure has been opportunely divided into different stand-alone modules that try to emulate several entities involved in a cellular transmission, guaranteeing the correct isolation of logic flows, and a simpler inter-modules communication.

The whole structure is depicted in figure 3.1.

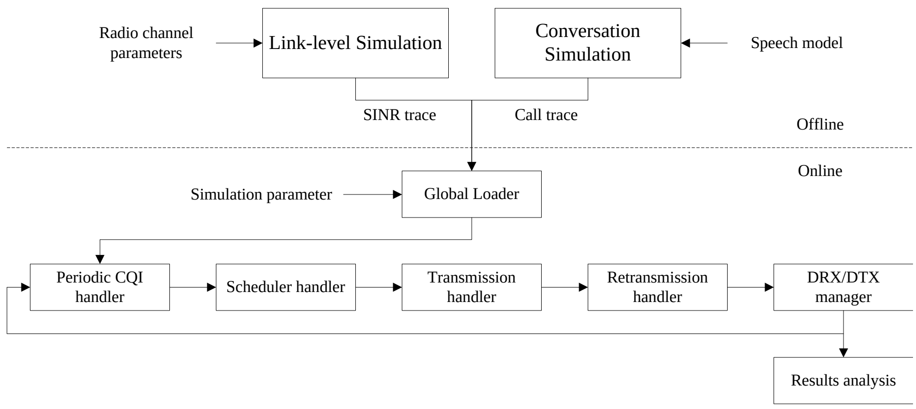


Figure 3.1: Simulator architecture

The main program is basically divided in two big sections: the first one initializes all needed variables and buffers, plus the voice call trace. The call dynamic is indeed defined *offline*, before the system level simulation, and it is an execution of a conversational speech model, obtained with a finite state machine, explained in the next section.

After the definition of the input buffer dynamics, the main loop of the system level simulation is then applied: all the required modules are called in order: channel quality reporting handler, scheduler, transmissions handler, retransmissions handler, DRX/DTX manager. The modularization of the script allows to call different kind of scheduler as well as channel quality report policies, with every possible combination. For each script cycle execution both downlink and uplink section of simulator run simultaneously.

We now introduce the entities running during the simulation, explaining them in details in the rest of chapter.

Channel quality feedback handler This module runs periodically, or it is *ad hoc* triggered in particular configurations and events combination. The purpose is to keep two memories for downlink and uplink of latest available values of channel quality feedbacks. The refresh of values can be done following the first or the third of the three policies explained in section 2.4.4. The simultaneous event of a CQI reporting and a transmission is forbidden to comply with [5].

Scheduler Schedulers implementations consist in four instances of scheduling entities dedicated to SPS/DS policies in DL/UL. The execution flow involves a simple VAD that detects the presence of new data in the input buffer. This event, in SPS module, triggers the scheduling systems in the particular *Talk Spurt Active* state, needed to handle the periodical allocation of resources, and their release. Another key feature linked to this state transition is the activation of short DRX/DTX condition. The scheduling trigger starts the analysis of channel quality feedback memory to perform the MCS selection, with the purpose of a 10% BLEP transmission. The successive step is the execution of time and frequency domain scheduling policies described in section 3.6.

Transmission The transmission module is a very simple probability analysis: to emulate the block error, we perform a calculation of BLEP value considering the effective channel conditions of PRBs involved in the transmission (i.e. PRBs allocated to the user in the current TTI). Then the success of failure of the transmission is modeled with the inverse percentile transformation method.

Retransmission The retransmissions handler is implemented with a pool of HARQ processes emulators, that can manage up to eight instances, as specified by 3GPP. The scheduling of retransmission data involves different calculation and actions compared to first transmissions one, but the transmission part is pretty similar. However, a simple model of incremental redundancy (IR) intervenes in the BLEP calculus. Retransmission handler is able to communicate with scheduler in SPS mode, to force an MCS renegotiation under particular circumstances.

DRX/DTX manager The activity state due to DRX/DTX events is tracked by this module, that provides interfaces to tune and reconfigure DRX parameters during the simulation, to trigger the UE in active state (emulating the power transition states delays too) and to check the instantaneous availability for transmission. This module provides a detailed log of instantaneous and average power consumptions, that we use to evaluate the different configurations under examination.

Next sections will illustrate and analyze the detailed functionalities implementations.

3.2 Conversational Speech: Brady Model

The first step for an accurate design and analysis of communication systems is to model the voice traffic. In VoIP traffic, particularly, it is important to understand how the speech process happens in a two-way phone conversation. This means that talk-spurts, silences, double-talks, and all the interaction between parties, are statistically relevant as speech patterns. Much research has gone into this area, from the simplest on-off states to a more complex models based on Markov Chains, but the final result establishes that Markov process can describe the conversational dynamic better than other [20].

It is also possible to differentiate each model considering their accuracy, i.e. the ability to predict the length of ten speech events [20]: talk-spurt, pause, double-talk, mutual silence, alternative silence, pause in isolation, solitary talk-spurt, interruption, speech after interruption, speech before interruption. In this scenario, the Six-State Markov Chain Model, also called Brady Model (figure 3.2), is of interest because of the ability to realize those aims [20], and for this reason we consider Brady Model useful to model and generate the VoIP traffic in our simulator.

The model considers speaker A's behaviour, governed by Poisson process [10], engaged in conversation with speaker B. There are six states A can be in, at any instant of time. The transitions among those states are controlled by the α parameters. Those parameters are not probabilities. However, if any α is multiplied by dt, $\alpha \cdot$ dt represents the probability that A will leave the state where it is in, at a given instant, to another one, during the next dt seconds: ´´transitional" probability [10]. Since A's behaviour can be see as a succession of Bernoulli trials [10], the transitions between two states, and so the correspondent α parameters, are functions only of the state where A is in, and not of the duration of occupation of a state. The only exception for this rule is actuated to guarantee silences $> 200ms$ and talk-spurt $> 15ms$, modifying some model's parameters to block the states transitions for at least 200 ms when A is in a silent state, and 20 ms when A starts to talk.

Once the conversation dynamics is defined, it is necessary to translate it into a bit stream in order to simulate the effective voice transmission process in a packet network system. 3GPP defines the Adaptive Multi-Rate speech codec as mandatory codec for 3G systems. For this reason we choose the AMR codec to generate voice frames in our simulator, following the codec parameters defined in [24] table 3.1.

In conclusion, the final traffic model considers a generic user A engaged in a two-way conversation with another user B. Every TTI A can be in one of the six states defined by the Brady Model. The AMR speech codec generates a voice frame of 320 bits every 20 ms and a SID frame of 120 bits every 160 ms after

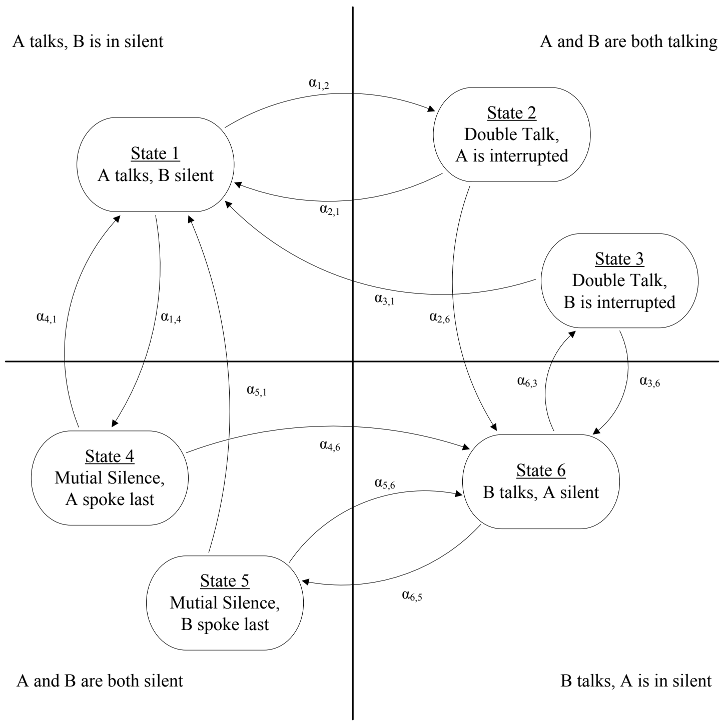


Figure 3.2: Brady model

Table 3.1: Main parameters of the traffic model

Parameter	Characterization
Codec	RTP AMR 12.2, Source rate 12.2 kbps
Encoder frame length	20 ms
SID payload	Modelled 15 bytes (5Bytes + header) SID packet every 160ms during silence
Protocol Overhead with compressed header	10 bit + padding (RTP-pre-header) 4Byte (RTP/UDP/IP) 2 Byte (RLC/security) 16 bits (CRC)
Total voice payload on air interface	40bytes (AMR 12.2)

the talk-spurt end, both for A user and B user. As concern the A's behaviour, it is handled as a succession of Bernoulli trials. So a one-step state transition matrix is defined in order to identify the allowed state transitions and their own probabilities, where the definition of the transitional probability is:

$$P\{S_i \rightarrow S_j\} = \alpha_{ij} \cdot dt \tag{3.1}$$

The state transition parameters are defined in Table 3.2:

Table 3.2: State transition parameters for the Brady model

$\alpha_{1,4}$ $\alpha_{6,5}$	$\alpha_{3,1}$ $\alpha_{2,6}$	$\alpha_{2,1}$ $\alpha_{3,6}$	$\alpha_{4,1}$ $\alpha_{5,6}$	$\alpha_{5,1}$ $\alpha_{4,6}$	$\alpha_{1,2}$ $\alpha_{6,3}$
0.7941	2.157	2.325	2.222	1.044	0.2785

In order to respect the condition the talk-spurt must be at least 20 ms and to follow the encoder frame length, the value of dt has been chosen to be 20 ms. The normalized transition probability matrix P therefore is:

$$P = \begin{pmatrix} 98,9274 & 0,2785 & 0 & 0,7941 & 0 & 0 \\ 2,3245 & 95,5183 & 0 & 0 & 0 & 2,1572 \\ 2,1572 & 0 & 95,5183 & 0 & 0 & 2,3245 \\ 2,2222 & 0 & 0 & 96,7340 & 0 & 1,0438 \\ 1,0438 & 0 & 0 & 0 & 96,7340 & 2,2222 \\ 0 & 0 & 0,2785 & 0 & 0,7941 & 98,9274 \end{pmatrix} \tag{3.2}$$

It is possible to identify the steady state probabilities vector π with:

$$\pi = \lim_{k\to\infty} [P]^k \tag{3.3}$$

that gives the following results:

Table 3.3: Steady state probabilities

P_1	P_2	P_3	P_4	P_5	P_6
38,31	2,38	2,38	9,31	9,31	38,31

and the expected permanence time in each of the six model's state:

$$E[T_j] = \frac{1}{1 - P_{jj}} \tag{3.4}$$

Table 3.4: Permanence time

P_{11}	P_{22}	P_{33}	P_{44}	P_{55}	P_{66}
1,8646s	0,4463s	0,4463s	0,6124s	0,6124s	1,8646s

The distributions of silence and active state occupancy duration and their expected permanence time are shown in figure 3.3.

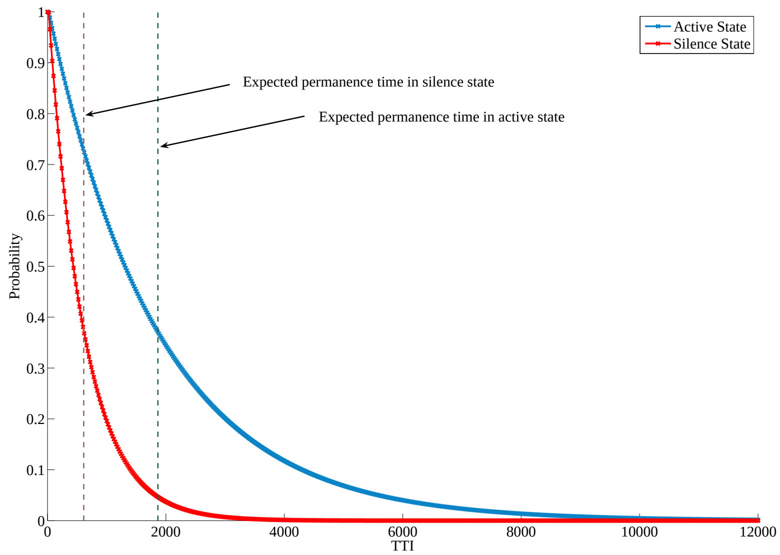


Figure 3.3: Brady model silence and active state occupancy duration

3.3 VoIP characterization

Once defined the voice activity model, it's indispensable to define how voice signal is digitalized and packeted to be fit to a packet switched network like LTE.

Adaptive Multirate (AMR) is an audio compression scheme realized for speech coding. It includes a multirate speech codec, a rate scheme controlled by source helped by a voice activity detector (VAD), a comfort noise generator and a set of functionalities for error masking. Available source rates are from 4.75kbps to 12.2kbps with a sampling rate of 8kHz. The rate setting can be adapted every 20ms, i.e. the period of packet generation. During silence periods, Silence Descriptor (SID) packets are generated with a periodicity of 160ms. The simulator uses a fixed 12.2kbps source rate, and applies ROHC that reduces protocol total overhead and compressed headers to the values indicated in table 3.1.

VoIP is a conversational class service: it bases its satisfaction criteria on mouth to ear voice delay induced by all network's interfaces. In fact, unlike full buffer traffic like web browsing, it needs real-time responsiveness and a guaranteed bit rate with tight delay constraints. The maximum mouth to ear delay acceptable without limiting the user interactivity is in the order of 150ms as said in chapter 1. The E-model rating graph (figure 1.1) classifies the users satisfaction based on the percentage of packets received with a certain amount of delay.

This delay budget must be divided among the segments of network located between sender and receiver equipments. The simulator analyzes the total air-interface delay including scheduling delay and first transmission plus HARQ retransmissions: all packets that take more than 50ms to be transferred both in downlink and uplink are discarded. In wireless communication systems, the loss of a small fraction of the data it's unavoidable. With VoIP traffic, the standards mark users as satisfied if not exceeding a 5% outage. For LTE, an outage of 2% of packets that exceed the delay budget during a call, is considered.

Another metric to evaluate the performance of the simulated system is the maximum amount of satisfied users loadable into it.

3.4 UE mobility: radio channel

After defining voice behaviour, it is necessary to focus the user in a possible radio channel experience. 3GPP establishes three different types of channel model for this purpose: Typical Urban, Rural Area and Hill Terrain channels [2].

Both for the downlink that the uplink, we choose Typical Urban as channel trace with 10 MHz of bandwidth and suppose the user is moving along the cell edge with a speed of $3km/h$. We choose the cell edge condition to assume the worst situation for an UE, in fact this is the cell part affected by the highest interference given by neighboring cells signals. The whole channel bandwidth is divided in 50 PRBs, but only in downlink is possible to allocate all the PRBs for the user dates. In uplink the 2 PRBs at the bandwidth edges are reserved for control transmissions.

In downlink, the G-factor is set to 0 dB, during the whole simulation. The G-factor is defined as:

$$G_{factor} = \frac{P_{eNB}}{P_I + P_N} \tag{3.5}$$

where the P_{eNB} is the eNodeB power, P_I is the cell interference power and P_N the power noise perceived by the UE. The figure 3.4 shows the channel selectivity.

For the uplink the user overall path gain is -125.65 dB. The trace gives information about the UE signal power on each 48 PRBs and relative interference. The figure 3.5 shows the final SINR trace for each PRB obtained as:

$$SINR_{PRB_i} = \frac{P_{S_i}}{P_{I_i}} \tag{3.6}$$

where $SINR_{PRB_i}$ the SINR for the i-th PRB in TTI j and P_{S_i} is the signal power of that PRB in TTI j and P_{I_i} the correspondent interference power.

Both the traces used in the simulator are obtained as a link level system simulation and are modeled with the parameters in Table 3.5

Since the downlink and uplink link-level traces are produced by two different simulators, we have checked that DL and UL signal properties are compatible with the assumed position of UE at the cell edge. For this reason we used figure 3.6 to map the uplink average SINR with the corresponding G-factor value. The red point shows the UE working point.

Figure 3.7 compares the average SINR of downlink and uplink trace. It is possible to note the faster variance of channel condition and lower values of uplink SINR.

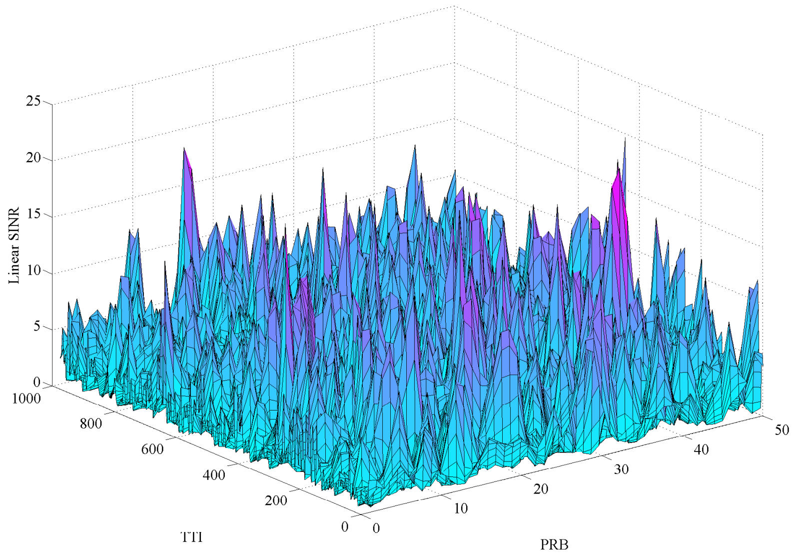


Figure 3.4: Downlink SINR trace

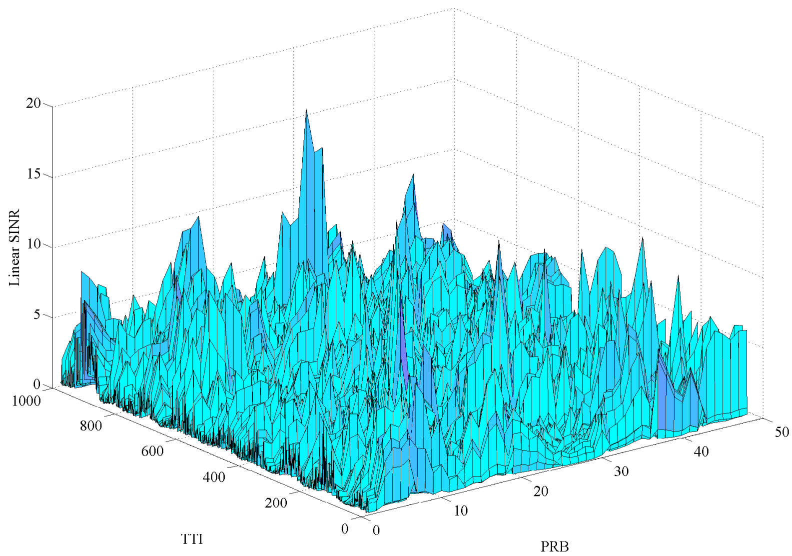


Figure 3.5: Uplink SINR trace

Table 3.5: Simulator traces parameters

Parameter		Value
Call time length		80s
Trace time length	Downlink	11000ms
	Uplink	1100ms
Antenna scheme		$1 \times 2(MRC)$
Temporal resolution		1ms (1TTI)
Bandwidth	Downlink	10MHz
	Uplink	10MHz
Frequency resolution	Downlink	1 sub-carrier (15kHz)
	Uplink	1 PRB (180kHz)

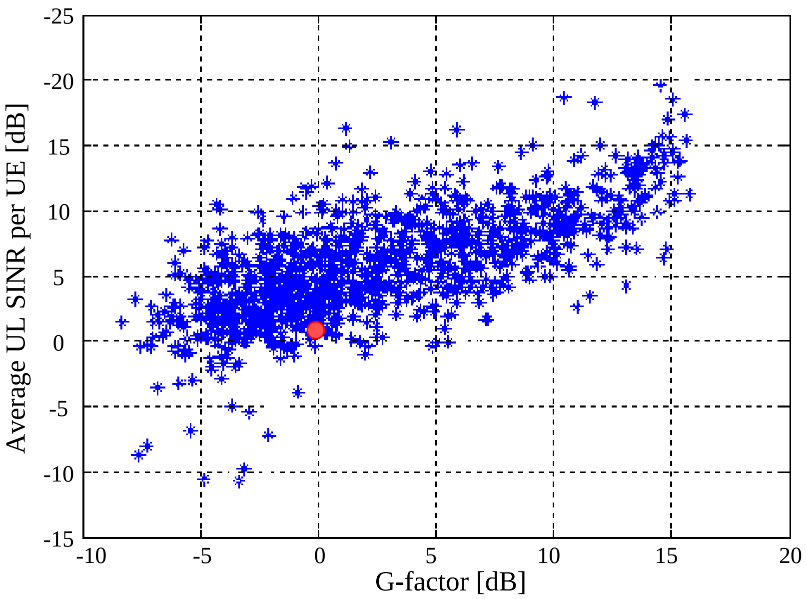


Figure 3.6: Downlink G-factor versus average UL SINR. Red point shows UE's working point

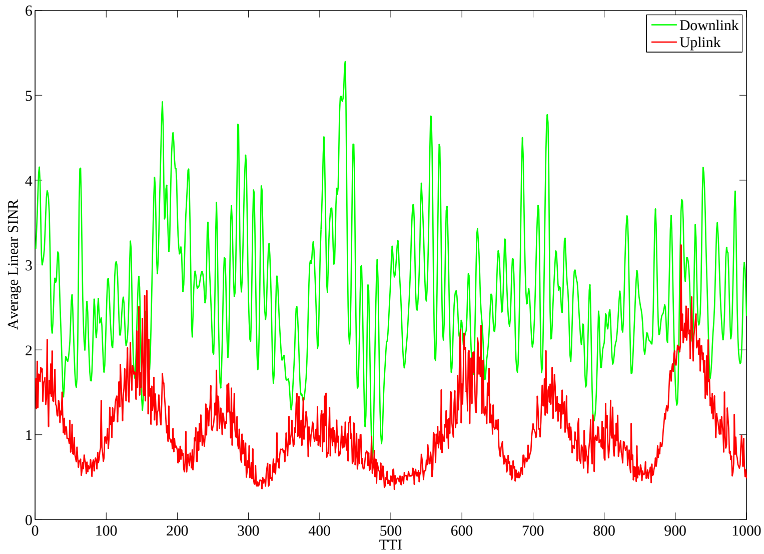


Figure 3.7: Downlink and uplink average SINR

3.5 Link Adaptation

As seen in section 2.4.5 the knowledge of the channel state is a prerogative for the eNodeB to handle the link adaptation in LTE. The model of link adaptation we propose and implemented in our simulator bases its decision on CQI and SRS reported by the UE. In downlink transmission, the CQI gives informations about the user channel condition as indication of the highest MCS that ensures $BLER \leq 10\%$. That indication is correlated to the user SINR as illustrated in figure 3.8.

So the reported MCS index allows to achieve the BLER target with the value of SINR evaluated in the measurement process as a geometric average of N contiguous PRBs, where N is the frequency resolution:

$$SINR_{AVG} = (\prod_{i}^{N} SINR_i(t))^{\frac{1}{N}} \tag{3.7}$$

The frequency resolution is set according to the chosen policy for time and frequency domain schedulers. The periodicity is instead related to the packet scheduling algorithm and can be selected by the eNodeB among the following types:

1. Periodic CQI reporting (both for semi-persistent and dynamic scheduling)
2. Periodic CQI reporting and Aperiodic CQI reporting required for MCS reconfiguration (only for semi-persistent scheduling)
3. Periodic CQI reporting and Aperiodic CQI reporting required at talk-spurt beginning (only for semi-persistent scheduling)

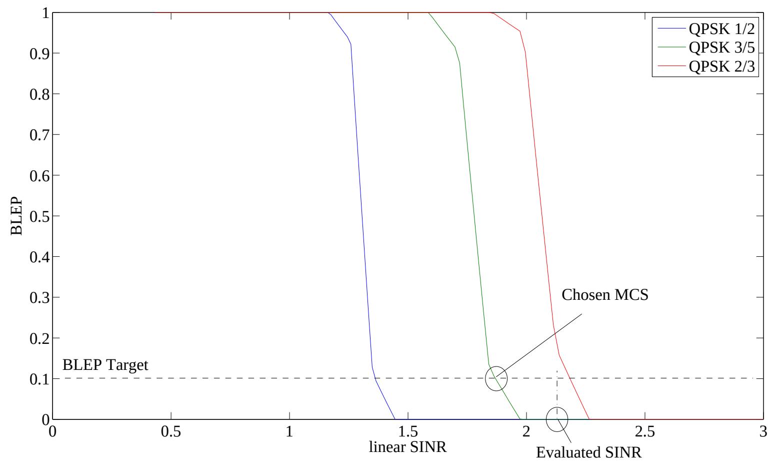


Figure 3.8: SINR versus BLER mapping function used to find the MCS that ensure BLER target

4. Periodic CQI reporting and Aperiodic CQI reporting required at talk-spurt beginning and for MCS reconfiguration (only for semi-persistent scheduling)

Those types have been proposed in order to evaluate the impact of different CQI reporting strategies on final BLEP and packets delays. In our study we will not consider CQI measurement errors but we will take into account reporting delays. Those delays are due to the time interval between the instant the CQI is evaluated in, and the instant when it is actually used by the eNodeB. The feedback inaccuracy affects the link adaptation causing a wrong MCS selection and consequently high BLER. For this reason it is mandatory to adopt some solutions to be able to achieve a $BLER \leq 10\%$ like Time and Frequency Aggressiveness and Outer Loop Link Adaptation, that will be introduced in section 3.6. The following scheme (figure 3.9) summarizes the link adaptation that is performed in downlink and uplink.

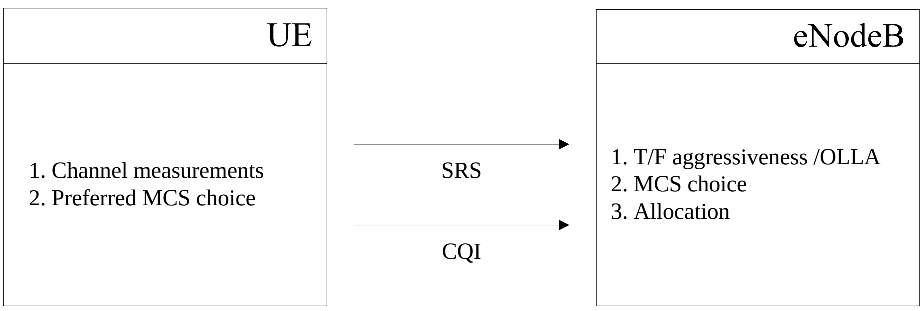


Figure 3.9: Link Adaptation Scheme

In uplink transmission, the eNodeB makes its own MCS evaluation according to the SRS sent by the UE. The selection process of the MCS index is the same of that made in downlink, but handled by the eNodeB. The evaluated SINR in this case is obtained considering the sounding signal value and the corresponding interference. Furthermore the eNodeB can choose the SRS periodicity between the following two types:

1. Periodic SRS reporting (both for semi-persistent and dynamic scheduling)
2. Periodic SRS reporting and Aperiodic CQI reporting required for SRS reconfiguration (only for semi-persistent scheduling)

The considered MCS and the corresponding values of PRB size for downlink and uplink are summarized in table 3.6

Table 3.6: MCSs for DL and UL

MCS	DL PRB size (bits)	UL PRB size (bits)
QPSK 1/10	Not used	30
QPSK 1/6	Not used	50
QPSK 1/5	59	Not used
QPSK 1/4	74	75
QPSK 1/3	98	99
QPSK 2/5	118	Not used
QPSK 1/2	147	149
QPSK 3/5	176	Not used
QPSK 2/3	196	199
QPSK 3/4	221	224
16QAM 2/5	236	Not used
16QAM 9/20	266	Not used
16QAM 1/2	296	296
16QAM 11/20	325	Not used
16QAM 3/5	355	Not used
16QAM 2/3	394	394
16QAM 3/4	443	443
16QAM 4/5	473	Not used
16QAM 5/6	493	493
64QAM 1/2	Not used	444
64QAM 3/5	532	Not used
64QAM 5/8	554	Not used
64QAM 2/3	591	592
64QAM 17/24	628	Not used
64QAM 3/4	665	666
64QAM 4/5	710	Not used
64QAM 5/6	739	740
64QAM 7/8	776	Not used
64QAM 9/10	798	Not used

3.6 Scheduling

The modeled scheduler entity, as described in section 2.4.2, firstly works in time domain to calculate a subset of all possible candidates to be scheduled in the next TTI. Then applies an allocation algorithm to guarantee fairness among users, and utilize all the resource units available in the best way.

The simulator supports two type of scheduling algorithms that exhibit different behaviors for both time and frequency domains, dynamic packet scheduler and talk-spurt based semi-persistent packet scheduler.

Both of them, in time domain, firstly choose among users scheduled for the next TTI: the choice is a combination of HARQ pending processes and normal transmissions. Although the simulator runs a single user per time, it is able to emulate other users requests, filling the transmission and retransmission queues pools with data obtained with other asynchronous simulations. The simulator scheduler checks if there are enough resources in current TTI to be able to transmit a single VoIP packet (assuming packet bundling and segmentation is not implemented).

Successive steps, in frequency domain, are emulated through an allocation algorithm that finds the best position for the needed PRBs. Scheduling activity is mainly based on a memory of channel quality indicators: the benefits from FDPS are maximized when the CQI reporting provides a reliable estimation that helps making a conservative predictive choice.

One of the main purpose of this report, is to investigate the actual pros and cons of two different approaches finding the best trade-off between network signaling load and scheduling accuracy. The following sections explore the actual implementation of dynamic and a semi-persistent scheduling in the simulator

3.6.1 Dynamic packet scheduling

The dynamic packet scheduler implementation extends the time domain selection of candidate users with a control that limits the cardinality of SCS to ten user, caused by the total amount of control channels and consequentially the amount of UEs the eNodeB can manage together in a single TTI.

The selection of best frequencies for each UE, i.e. candidates PRBs, is based on a ranking function that exploits the CQI granularity over the whole bandwidth. The minimum value of CQI granularity of 3 PRBs is adopted. The implementation is a simple stack of 3 adjacent PRBs groups #IDs, orderer by latest available value of CQI, popped each time other PRBs are needed or the current ones are already allocated. The total amount of needed PRBs is evaluated by the simulator to fit a VoIP packet with the MCS reported by link adaptation algorithm. If the number of needed PRBs exceeds the available ones in the current TTI, the schedulers skips it by one TTI.

The scheduler implementation deals with the DRX/DTX control entity too. When a new talk-spurt activity is detected, i.e. when new packets are queued into the transmission buffer, it triggers the short DRX cycle, to ensure a coherent periodicity and synchronization of DRX/DTX active states and codec packets. The return to normal long DRX/DTX cycle is automatically performed when transmission buffers and HARQ processes buffers are empty for a relevant amount of time.

Outer loop link adaptation analysis

To guarantee high transmission quality and BLEP target, the adopted solution under dynamic scheduling is outer loop link adaptation (OLLA). As described in 2.4.6, OLLA aims to adapt the MCS to the actual channel condition, processing the information provided by the CQI reporting. Since all the analysis are made considering linear SINR, the equations (2.2) and (2.3) have been modified to (3.9) and (3.8):

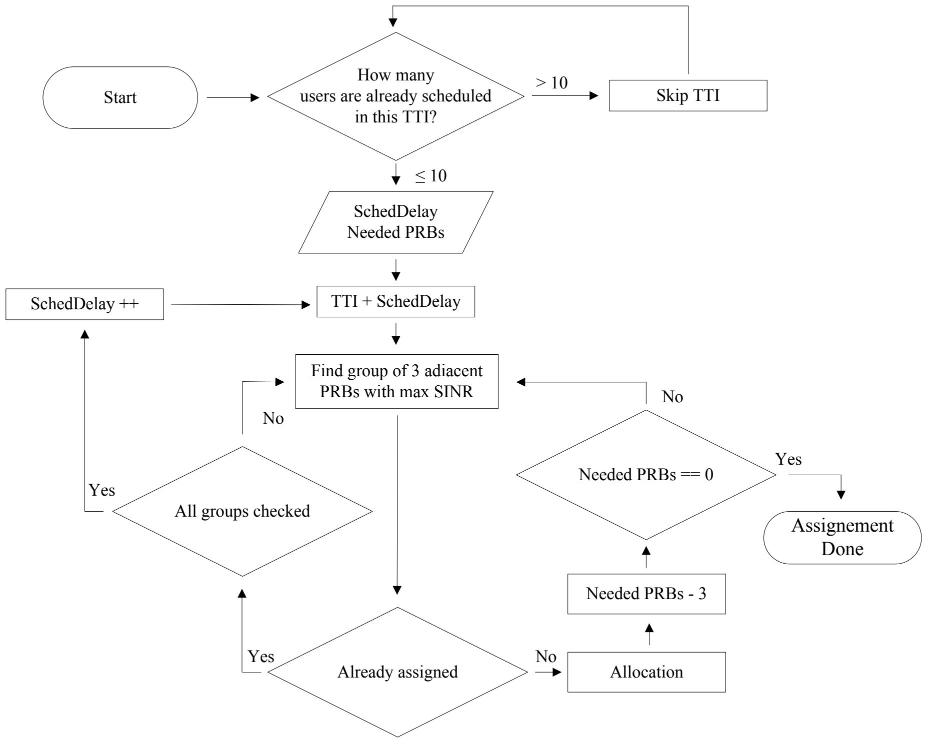


Figure 3.10: Dynamic Scheduling flow chart

$$A_{off} = A_{off} * A_{up} \tag{3.8}$$

$$A_{off} = \frac{A_{off}}{A_{up}} \tag{3.9}$$

OLLA works as a digital discrete controller whit A_{up} and A_{dw} control variables and their proper values must be identified in order to guarantee the target tracking. The following analysis are made considering a single user at the cell edge with channel state conditions described in 3.4. We will consider the effect of different value of A_{up} and A_{dw} on BLEP for growing values of CQI reporting period. Table 3.7 sums up the simulation parameters.

Table 3.7: Simulation parameters Sim #3.1

Simulation parameter	Value
Simulation Time	80 s
CQI granularity	3 PRBs
Bandwidth	10 MHz
BLER target	10%
A_{up}	$1(0dB), 1.0116(0.05dB), 1.0233(0.1dB),$ $1.1220(0.5dB), 1, 2589(1dB)$
CQI Reporting	$[5, 10, 15, 20, 25]$ms
SRS Reporting	$[5, 10, 15, 20, 25]$ms

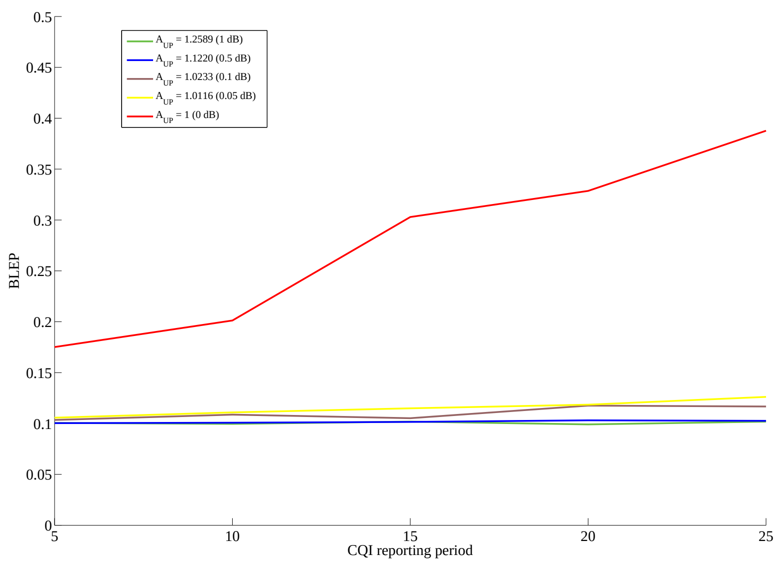


Figure 3.11: DS - BLEP trend in downlink for different A_{up} values and different CQI reporting periods

Figure 3.11 compares the downlink final BLEP trend using different values of A_{up} with the one with OLLA disabled. The convergence to the BLEP target is almost guaranteed for each value of A_{up} and different periods of CQI reporting. The delay in CQI reporting affects the precision of the MCS selection because of the basic inaccuracy of the channel condition reports. That is clearly shown in the case without correction.

In case of long CQI reporting periods (in order of 20/25 ms), an increased control action A_{off} can still provide the target tracking. However, higher values of A_{off} result in a too fast and very less accurate control action: if the CQI reporting delay is high (> 25 ms), the control can start to oscillate around the optimal value, becoming unable to stabilize itself, and finally increasing the overall BLEP.

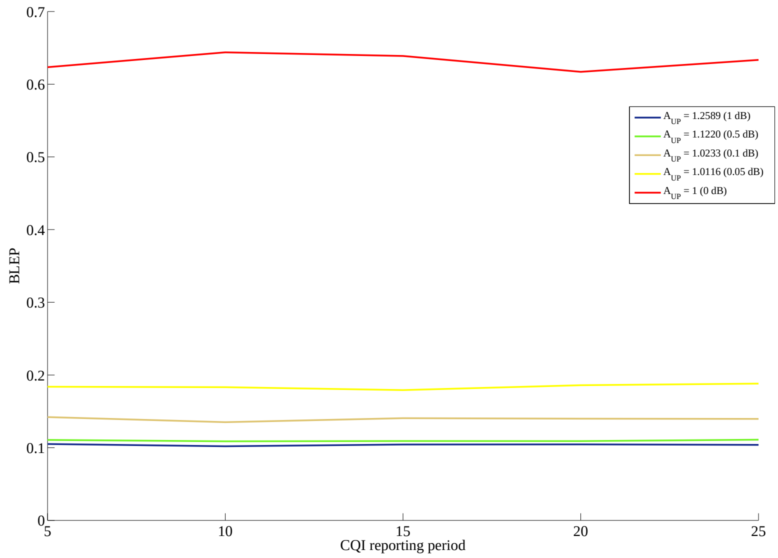


Figure 3.12: DS - BLEP trend in uplink for different A_{up} values and different CQI reporting periods

The results concerning uplink statistics are shown in figure 3.12. The values of BLEP are higher than downlink case because of the higher variance of the uplink SINR trace (see figure 3.7 in section 3.4), but a step of 1.1220 (0.5 dB) or 1.2589 (1 dB) is still able to make the controller to track the BLEP target. Therefore, for the following simulation we will consider $A_{up} = 1.1220$ (0.5) dB and $A_{dw} = 1.0129$ (0.0556 dB).

It's interesting to review how different values of OLLA offset change the control behavior and how it is reflected over the average number of per user assigned resources.

As explained before, high values of A_{off} can assure the BLEP target by means of a big control action. That means the perceived SINR is significantly reduced and consequently the link adaptation provides a lower MCS to guarantee reliable transmissions. The final result is an increased number of assigned

PRBs (figure 3.13 and 3.14).

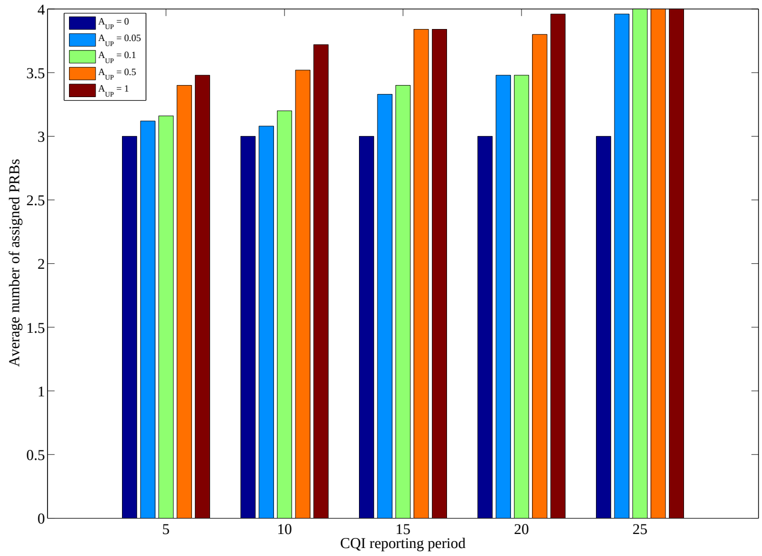


Figure 3.13: Downlink average number of assigned PRBs with OLLA offset and channel quality reporting period variation

3.6.2 Semi-persistent Packet Scheduling

The time domain scheduling implementation in semi-persistent scheduling is slightly different from dynamic one. The considered SPS modality is talk-spurt based, so a sort of VAD functionality helps to distinguish the talked parts of the call trace both in downlink and uplink. A the beginning of every new detected talk-spurt, the scheduler triggers a periodic (with appropriate periodicity) allocation of PRBs and a simultaneous message to DRX/DTX entity. The fixed allocation of resources in a talk spurt can still be forcibly released if to many retransmission events occurs. This is due to an evident bad MCS selection performed in a transitive good quality channel condition that results in an overestimation. In this case, a special procedure, we address as *MCS reconfiguration* is triggered: if eight first transmissions suffer from failures, the scheduler releases the allocated PRBs and restarts the MCS selection algorithm, taking care to adopt an higher order modulation or coding scheme. This procedure can be performed channel blindly too, if no CQI report are requested, only increasing the coding scheme.

The CQI reporting mode, chosen for SPS simulations, is Wideband CQI, so SPS can not exploit frequency diversity. The actual attempt is based on the statistical analysis of the common channel fading dynamics among the considered bandwidth. In fact, as seen in figure 2.10, there are lots of groups of adjacent PRBs that share poor channel conditions: the statistics suggests that the choice of non adjacent PRBs over the full bandwidth increases the chance

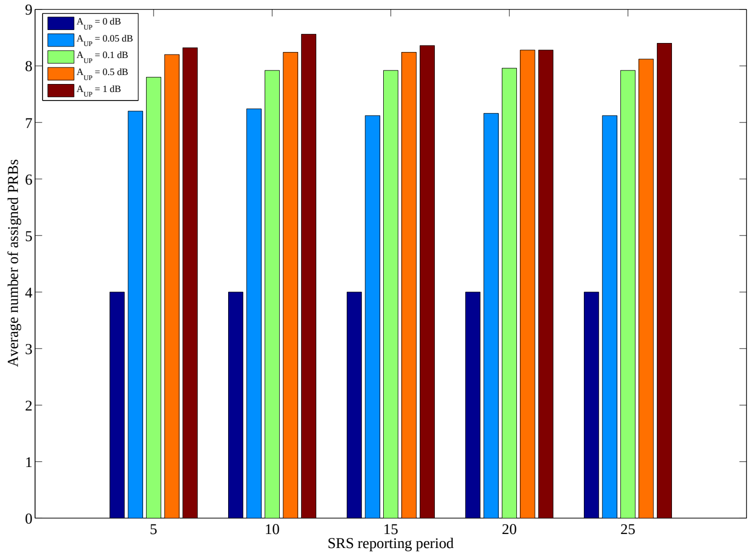


Figure 3.14: Uplink average number of assigned PRBs with OLLA offset and channel quality reporting period variation

to find high SINR portion of that, exploiting the maximum frequency diversity, when detailed informations of quality are not available [9]. The algorithm can be configured to use a fixed PRBs distance or calculate it dynamically: in the latter modality it makes use of values of allocated bandwidth, number of needed PRBs, and the allocation table for the concerned TTI to try to find a possible allocation, recursively skipping PRBs already assigned to other users and reducing PRBs spacing to avoid fragmentation of the set of available ones. Frequency hopping for SPS allocations is not implemented.

As specified in [5] SPS allocated resources are implicitly released when a configurable amount of them are not actually used. The release of resources coincide with the end of shortDRX condition too.

Time and Frequency Aggressiveness analysis

As previously explained, the SPS needs a conservative MCS choice to cope with the channel condition variations. The adopted solution provides to lower the user perceived SINR. Since the eNodeB receives channel state informations as MCS index, equation (2.4) relates the SINR value to be processed in order to choose a new more conservative MCS. The first processing step is a low pass filtering of the SINR reported by the UE in time domain:

$$SINR^{i}_{filtered} = SINR^{i-1}_{filtered} \cdot (1 - \alpha) + SINR^{i}_{cqi} \cdot \alpha \qquad (3.10)$$

where $SINR^{i}_{cqi}$ is the reported SINR, $SINR^{i-1}_{filtered}$ is the previous low pass filtered simple, and α the smoothing factor that we will refer to as Time Aggressiveness from now on.

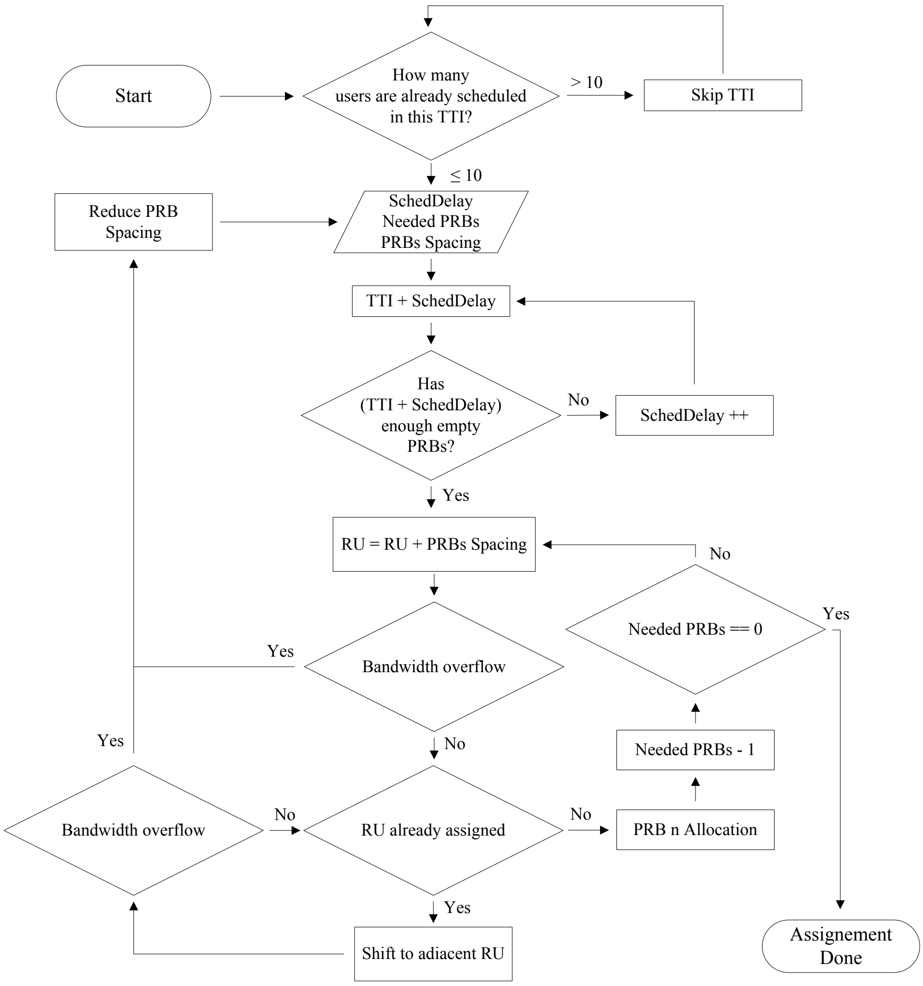


Figure 3.15: Semi-persistent Scheduling flow chart

The second step is a frequency reduction of $SINR^{i}_{filtered}$:

$$SINR_{proc} = SINR^{i}_{filtered} \cdot \beta \tag{3.11}$$

where β is a reduction factor we will call Frequency Aggressiveness.

The values of Frequency and Time aggressiveness can be chosen in [0,1]. So an accurate analysis is needed to find the best values in order to achieve the BLER target of 10 %. The simulation assumption are shown in table 3.8.

Table 3.8: Simulation parameters Sim #3.2

Simulation parameter	Value
Simulation Time	80 s
CQI reporting scheme	Wideband CQI
Bandwidth	10 MHz
BLER target	10%
Time Aggressiveness	$[0,9 : -0.1 : 0,6]$
Frequency Aggressiveness	$[1 : -0,05 : 0,6]$
CQI Reporting	60ms
SRS Reporting	20ms

We consider a single user at the cell edge with a channel state for downlink and uplink described in 3.4. The eNodeB requires a periodic CQI and SRS reporting respectively every 60 ms and 20 ms; an aperiodic CQI is required for talk-spurt beginning and MCS reconfiguration and an aperiodic SRS reporting for MCS reconfiguration.

All the possible combination are made considering value of Time aggressiveness between 0.9 and 0.6 with 0.1 of resolution and Frequency Aggressiveness between 0.1 and 0.6 with 0.05 of resolution. Each simulation is repeated 10 times.

Figure from 3.16 to 3.23 show the BLEP trend for each combination, in downlink and uplink transmission. We can see the BLEP is reduced lowering the frequency aggressiveness and the value of 0.65 seems to be the maximum value to guarantee a BLEP around the target of 10 % for downlink transmission. In uplink, instead, no values of frequency aggressiveness can ensure the convergence to the target. Nevertheless, we can allow a transmission with a BLEP around the target only choosing 0.6, or lower, as frequency aggressiveness. As concern the time aggressiveness, it can reduce the variance of BLEP values. That allow to increase the probability to have a BLEP closer to the target with less varibility.

The frequency aggressiveness effect is an under estimation of the channel state that provides a more conservative MCS choice at the beginning of the talk-spurt, as shown in figure 3.24, in order to cope with lower channel condition during the talk-spurt. In fact, we can see that a processed WB-SINR higher than the one perceived by the UE causes an higher BLEP in transmission.

At the same time we can not use too low values of frequency aggressiveness because that means a lower MCS selection and so a higher number of allocated PRBs for the UE during the talk-spurt (from figure 3.25 to 3.28).

As we can see in figures from 3.25 to 3.28 the bottleneck is the uplink. Considering 10 MHz of bandwidth there are 48 allocable PRBs and if the eNodeB sets the frequency aggressiveness to 0.6 the allocated PRBs are 8 for one UE,

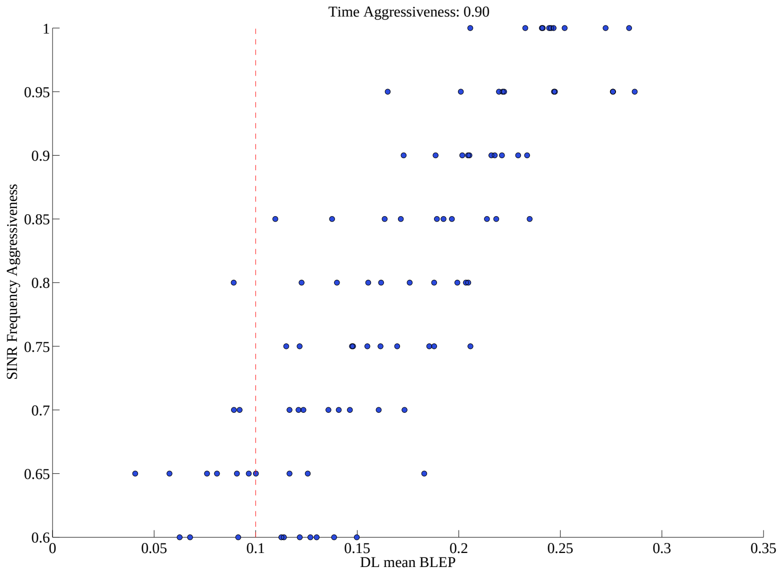


Figure 3.16: Downlink BLEP trend with Time Aggressiveness = 0.9

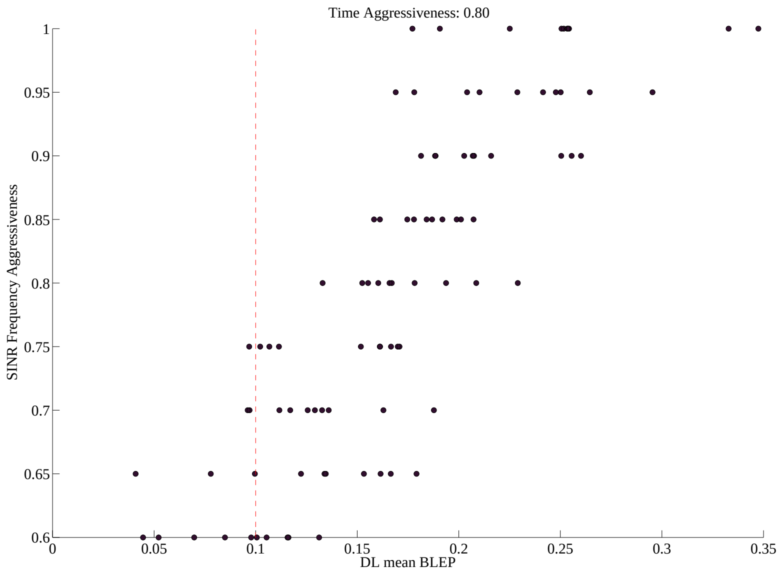


Figure 3.17: Downlink BLEP trend with Time Aggressiveness = 0.8

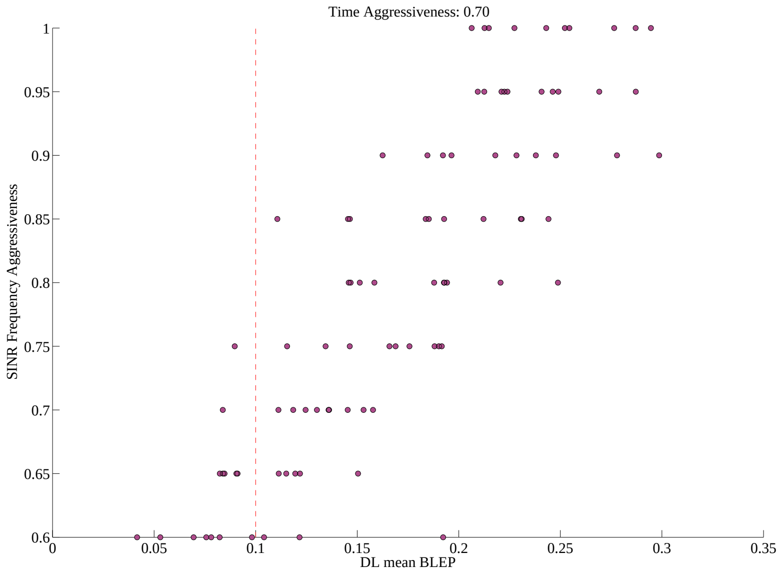


Figure 3.18: Downlink BLEP trend with Time Aggressiveness = 0.7

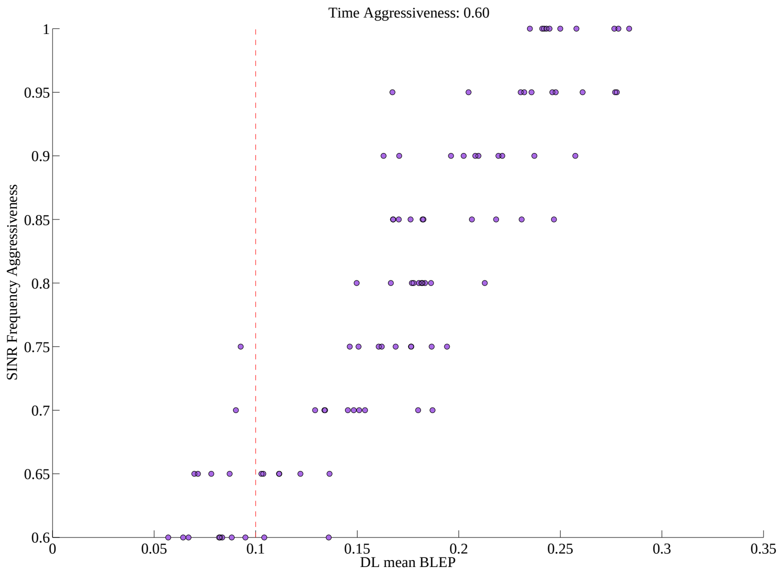


Figure 3.19: Downlink BLEP trend with Time Aggressiveness = 0.6

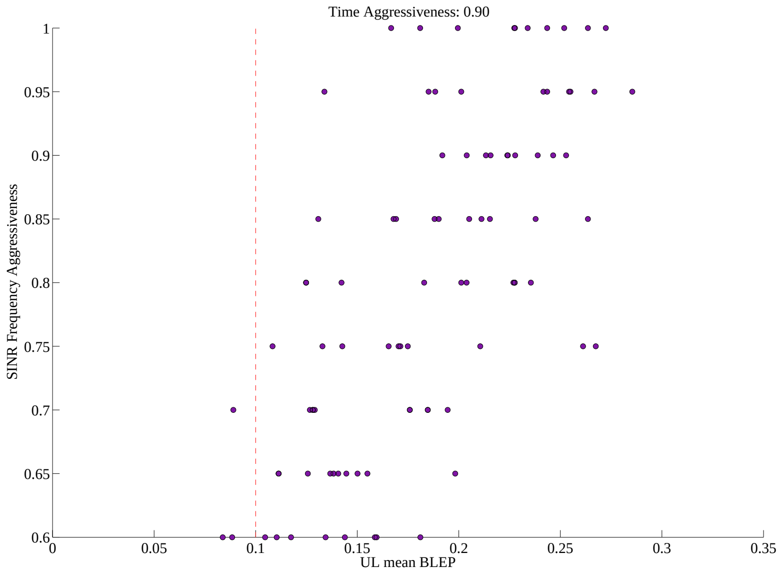


Figure 3.20: Uplink BLEP trend with Time Aggressiveness = 0.9

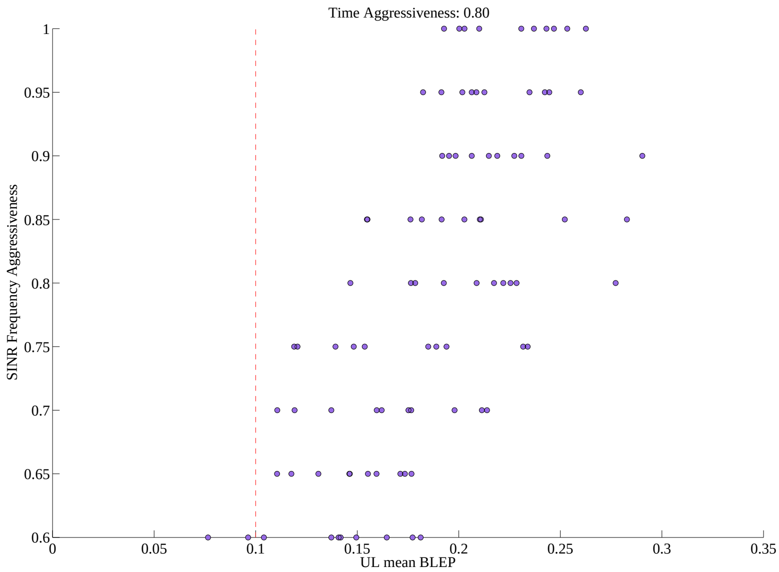


Figure 3.21: Uplink BLEP trend with Time Aggressiveness = 0.8

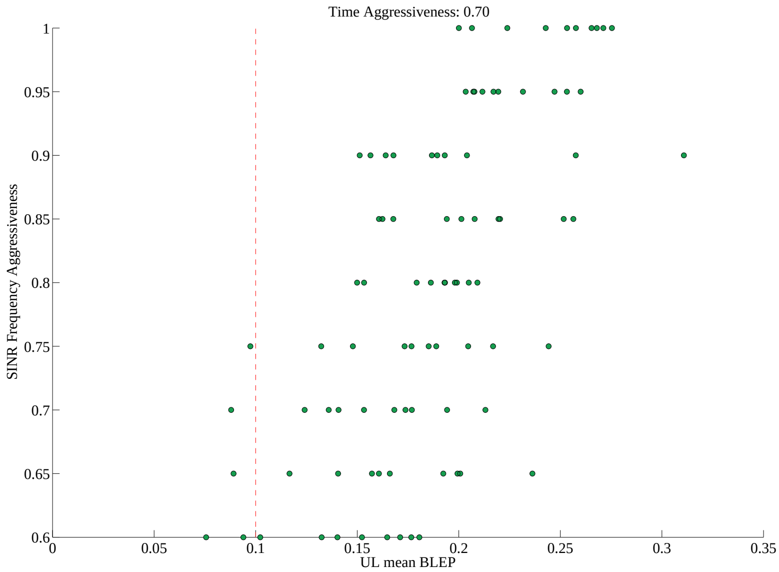


Figure 3.22: Uplink BLEP trend with Time Aggressiveness = 0.7

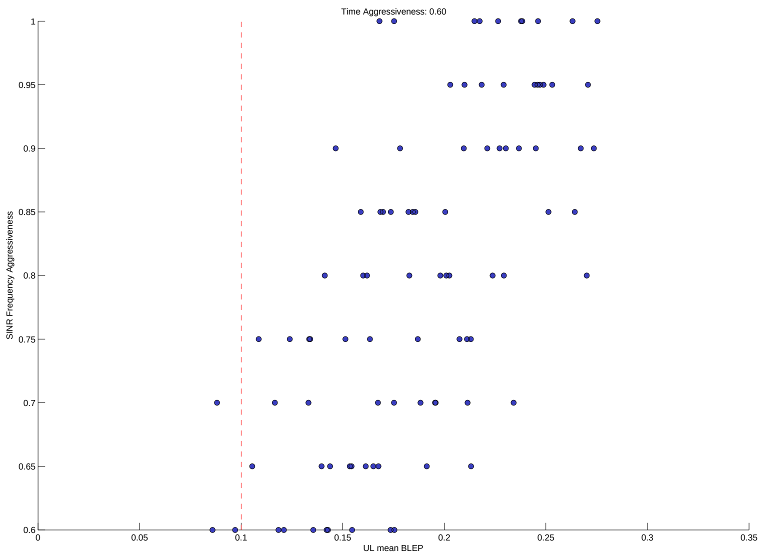
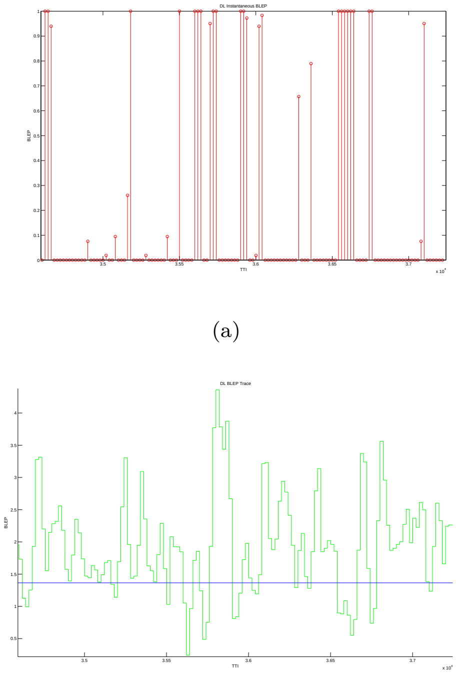


Figure 3.23: Uplink BLEP trend with Time Aggressiveness = 0.6



(a)

(b)

Figure 3.24: a) BLEP trace during talk-spurt in a generic call; b) Real WB-SINR versus estimated WB-SINR in the same talk-spurt

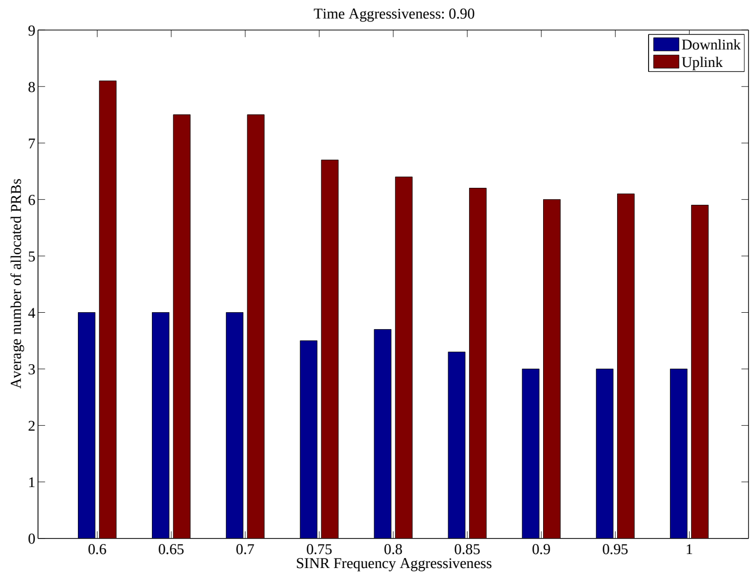


Figure 3.25: SPS - Average number of allocated PRBs in downlink and uplink transmission, Time Aggressiveness = 0.9

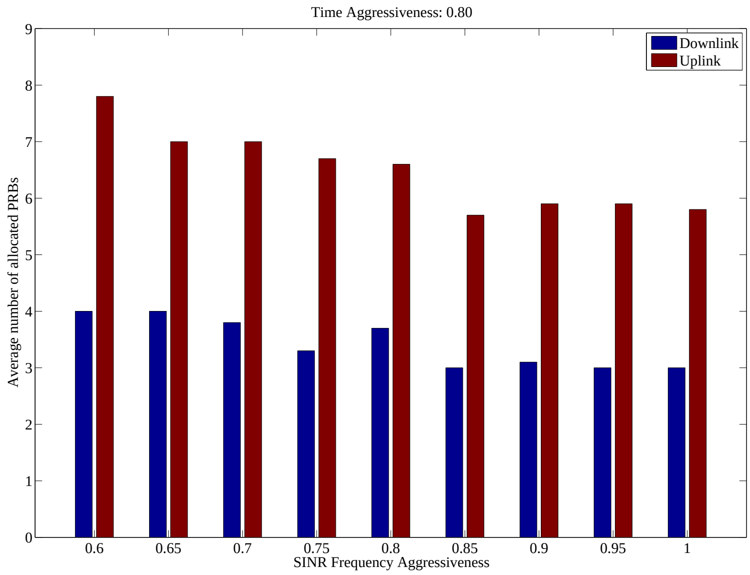


Figure 3.26: SPS - Average number of allocated PRBs in downlink and uplink transmission, Time Aggressiveness = 0.8

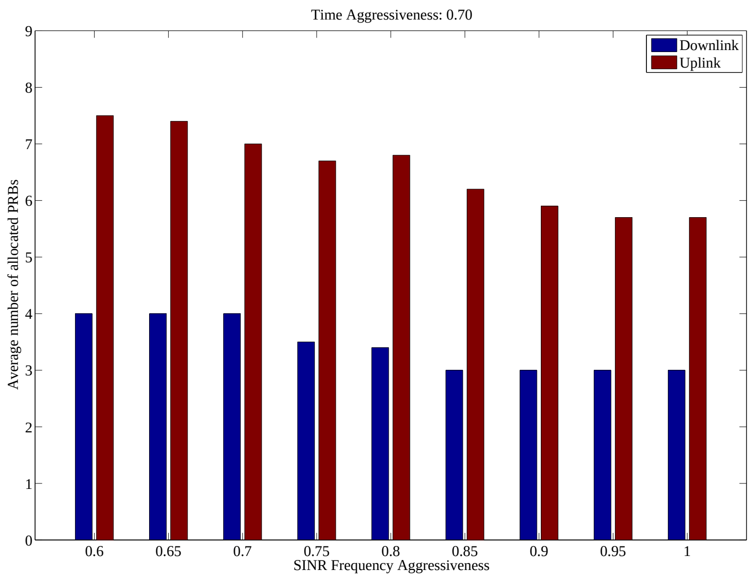


Figure 3.27: SPS - Average number of allocated PRBs in downlink and uplink transmission, Time Aggressiveness = 0.7

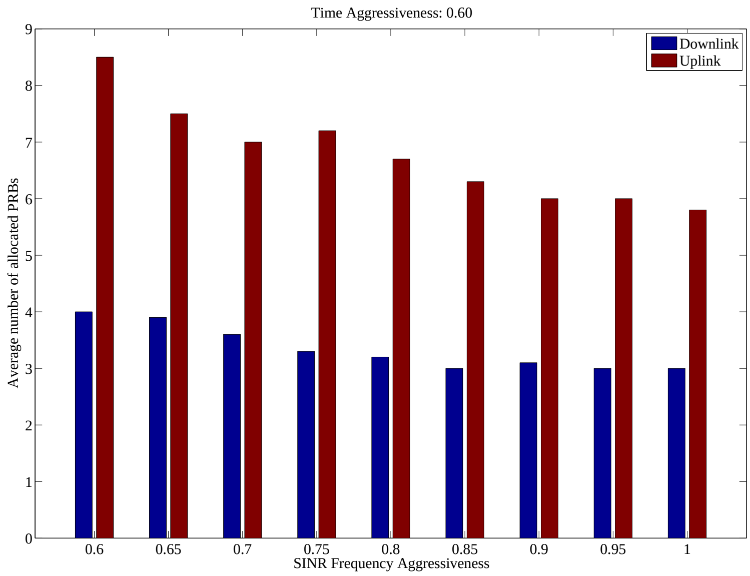


Figure 3.28: SPS - Average number of allocated PRBs in downlink and uplink transmission, Time Aggressiveness = 0.6

so the 16% of the resources are used only by that user in a TTI. If we consider 20 ms, i.e. the vocal frames generation interval for AMR codec, the number of schedulable users at the cell edge, talking at the same time, is 120. Obviously it is an ideal case because it is necessary to consider also other factors like other different traffic types, dynamically scheduled users etc., so the number of users at the cell edge that can be scheduled with semi-persistent scheduling should be much lower than 120 in 20 ms.

3.7 Transmission and HARQ Retransmission

In order to estimate downlink transmission quality and performance we choose Exponential Effective SINR Metric (EESM) to find the instantaneous channel condition as an effective SINR value. The Actual Value Interface (AVI) is used to obtain the received BLER as a function of the effective SINR and the MCS chosen by means of the link adaptation (figure 3.29).

The EESM model is represented by equation (3.12).

$$EESM = -\beta * \log(\sum_{i}^{N} e^{\frac{-\gamma_i}{\beta}}) \tag{3.12}$$

where β is a parameter obtained from link-level simulations for each adopted MCS, γ is the SINR of the i-th sub-carrier and N is the number of sub-carriers allocated for the UE.

In uplink transmission, it is not possible to use EESM mapping because of the different digital modulation scheme, based on SC-FDMA instead of OFDMA.

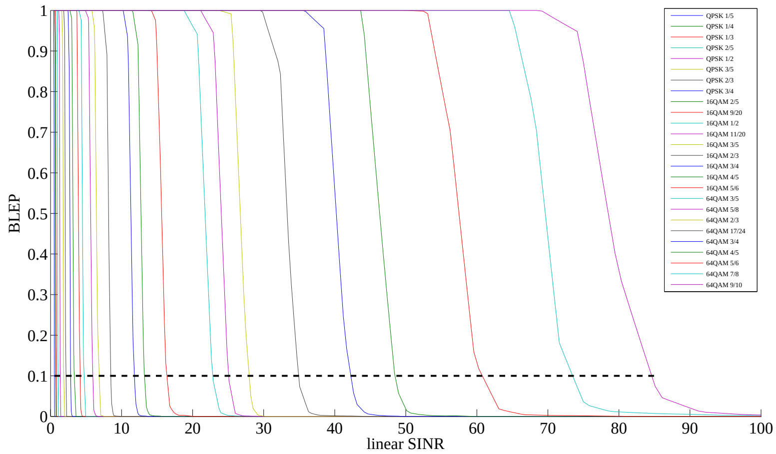


Figure 3.29: Downlink BLEP AVI curves

An alternative mapping scheme that involves a Maximum Ratio Combining (MRC) receiver can be applied to map the channel conditions into a single value of SINR. This algorithm is based on the knowledge of both signal and total interference powers the UE perceives on individual portions of bandwidth: the MRC functionality is represented by equation (3.13).

$$SINR = \frac{\sum_i^N S_i}{\sum_i^N I_i} \tag{3.13}$$

where S_i is the i-th PRB signal and I_i the relative interference, and N is the number of PRBs assigned to the UE. Like in downlink that value maps the correspondent BLER using the AVI approach (figure 3.30).

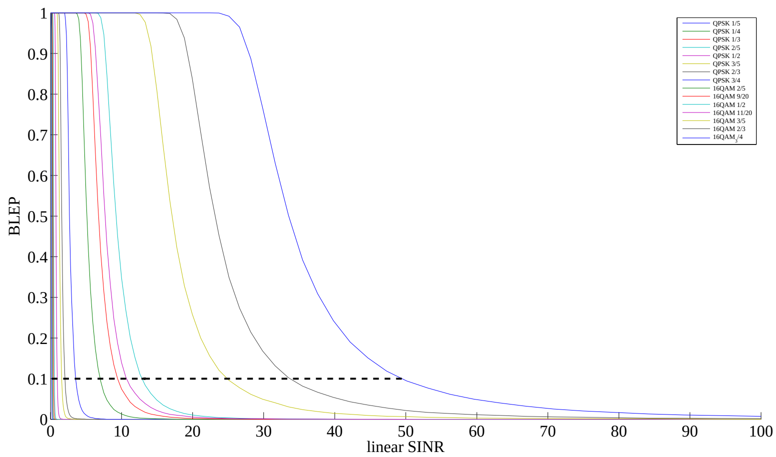


Figure 3.30: Uplink BLEP AVI curves

The final BLER gives a probability of transmission success or failure.

The acknowledgment for the packet at the TTI n is received after 4 TTI from the transmission. In case of NACK the packet is stored in the HARQ soft buffer and the HARQ transmitter entity will handle the packet retransmission. The processing time is set to 4 ms. So the UE can be scheduled at the TTI n + 8 (HARQ round trip time), both for downlink and uplink (figure 3.31).

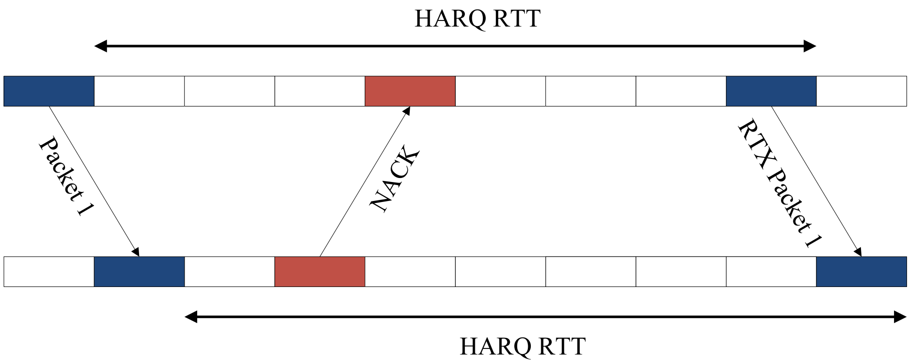


Figure 3.31: HARQ: The user can be scheduled at least after 8 ms

In downlink we adopt asynchronous adaptive scheme whereas synchronous adaptive scheme in uplink. Hence, both need a new link adaptation and a new MCS, but in downlink the packets can be scheduled as soon as possible (figure 3.32), in uplink, instead, there are predefined time instant to send the retransmissions.

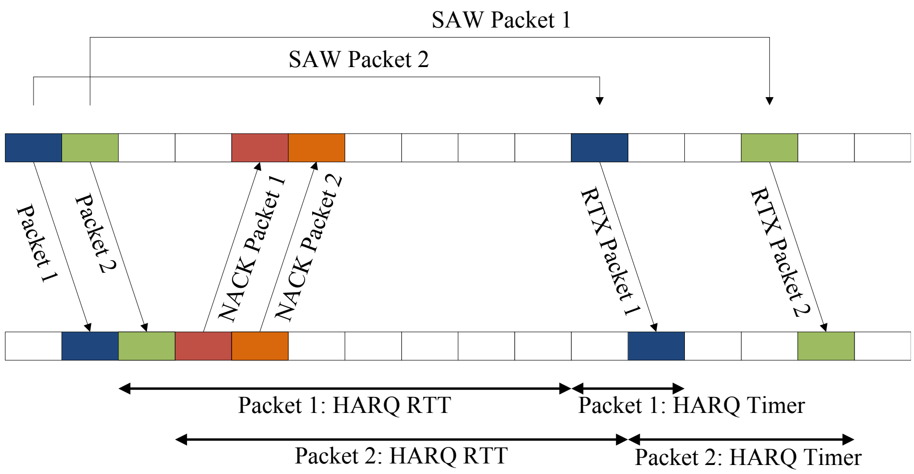


Figure 3.32: Downlink HARQ: the packet retransmission can be scheduled at least after 8 ms

After each retransmission the HARQ receiver entity combines the old packet with the new one and an improvement of the BLER is expected.

Because of the adaptive HARQ, a new MCS is chosen every retransmission considering the channel state. The model we adopt to improve the expected

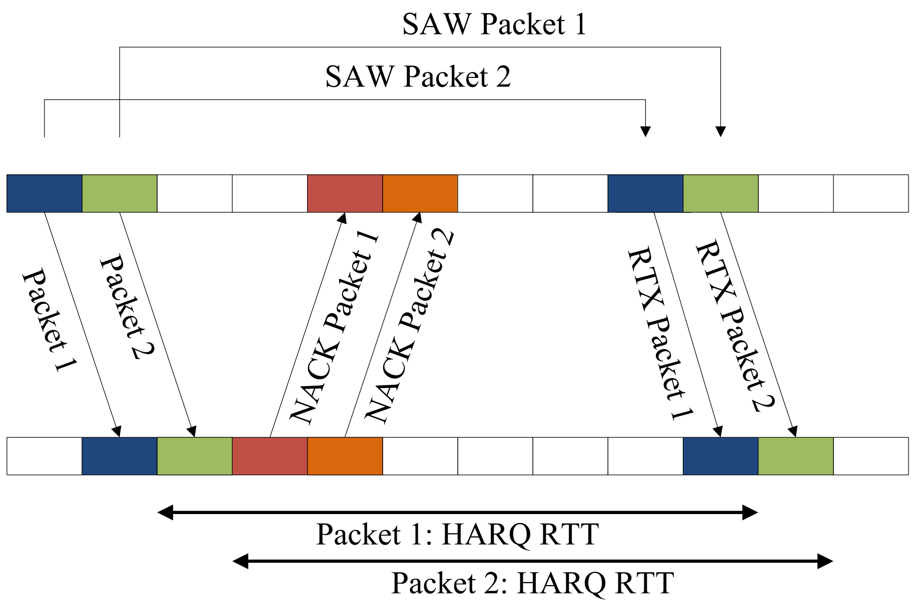


Figure 3.33: Uplink HARQ: the packet retransmission can be scheduled only after 8 ms

BLER is based on a simple recursive SINR sum:

$$SINR_{HARQ} = \sum_{i}^{N} SINR_i \tag{3.14}$$

where the $SINR_{HARQ}$ is the combined SINR after N transmission of the same packet. Since we consider IR as our combining strategy, each retransmission provides the receiver extra information, improving the final BLER. Hence the value of $SINR_{N-1}$ is a gain for the actual $SINR_N$ perceived by the UE and allows to simulate a shift of the SINR vs. BLER curves (figure 3.34).

We also set to 4 the maximum number of allowed transmission for each packet after which the packet is considered lost, as defined in LTE.

3.8 Power consumption

The UE power consumption depends by different components like active processes, display, applications, amplifiers, etc. Because of the impossibility to simulate all those factors we adopt a simplified model referred only to the RF modem power consumption based on the one described in [22]. In our study we need to consider both downlink and uplink behaviour. Hence, the model represented in figure 3.35 is modified with uplink power parameters from internal NSN specification [23].

The model identifies three main activity states: Deep Sleep, Light Sleep, Active, and include also fixed state transition times and relative power consumption. Those transition states are necessary to describe the circuit activation time between off and active state and are aligned to the simulator resolution

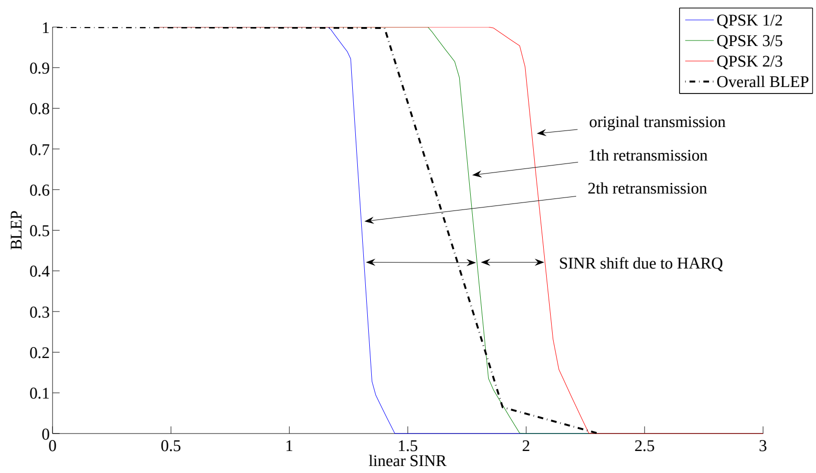


Figure 3.34: Overall BLEP after two retransmissions

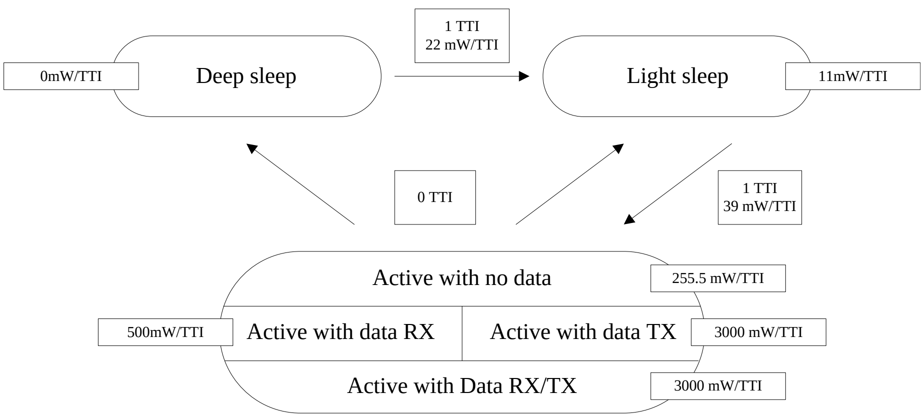


Figure 3.35: Power Consumption Model

set to 1 TTI. In active state it is possible to define three UE behaviour related to data reception, data transmission and control channel decoding. The value of power consumption for each main state are reported in [17] as a arbitrary average values.

In active state with data receiving we take into account packet reception, acknowledgment reception, and scheduling request response from the eNodeB. As concern the uplink behaviour, the active state considers data transmission, acknowledgment response, scheduling grant request, CQI and SRS report send by the UE. The power consumption value is not defined in [17] hence we arbitrary choose 3000 mW.

The information described above are utilized to calculate a power consumption average of the UE during the whole simulation.

Chapter 4

DRX/DTX with SPS

As previously explained, the target of this study is to analyze DRX/DTX strategies and performance for VoIP traffic in LTE, under two different packets scheduling algorithms. In the following chapter we start analyzing the UE performance under semi-persistent scheduling. This scheduling algorithm seems to be able to easily handle the allocation of periodic network traffic like VoIP. At the same time, because of the resource preallocation, the SPS can not adapt each transmission to the channel condition state. For this reason the first following analysis will be carried out through the SPS performance evaluation in terms of BLER target tracking, both for downlink and uplink. Afterwards, the analysis will be focused on the DRX/DTX functionalities performance to verify the impact on power consumption saving and packets delays. The purpose is to find out a set of parameters able to guarantee the best trade off between battery saving and user satisfaction.

The analysis will be lead by means of the Matlab®simulation tool we developed for the aim. Both user uplink and downlink behaviour will be carefully analyzed.

4.1 Modeling Assumptions

The simulations will be carried out from the point of view of a user that is moving along the cell edge with a velocity of 3 km/h, engaged in a bidirectional call with another user, whose call dynamics is defined by the Brady Model (section 3.2). Link adaptation and packet scheduling are implemented in detail as well as the UE DRX/DTX functionalities. The channel radio conditions are simplified assuming a single cell scenario as described in section 3.4: a SINR trace with Gaussian G-factor distribution set to 0 dB describes the UE channel experience in downlink, and a signal power trace and relative interference describes the one for the uplink.

The packet scheduler can handle different users in time and frequency domain, allocating resources among them and trying to ensure a transmission BLER below 10%, both for downlink and uplink. The scheduler must also take into account the UE's DRX/DTX behaviour in order to syncronize the user scheduling with DRX/DTX active states.

Uplink scheduling request delay, processing delay, air interface delay and network jitter are considered, respectively of 14 ms, 3 ms, 1 ms and [1 : 3] ms.

For the power consumption we will refer to the model described in section 3.8. Since the values reported in section 3.8 are indicatives, we can not consider the final power consumption value as reliable but it gives an important indication of the power consumption trend under the scenarios we are going to analyze.

The HARQ entity can handle max 8 stop-and-wait processes. Every retransmissions can be scheduled after 8 ms the first packet transmission, as described in section 3.4. The max number of consecutive transmissions for a packet is set to 4 and it is possible to perform a MCS reconfiguration after 8 first packet consecutive transmissions failures, also without CQI or SRS reporting request.

The system performance will be evaluated in term of power consumption and user satisfaction. i.e. the number of dropped packets due to a transmission delay higher than 50 ms.

4.2 SPS performance without DRX/DTX

We start analyzing a single and multiple users scenario without considering DRX/DTX strategies to verify the SPS performance in presence of a generic user at the cell edge. The proposed solution has been shown in section 3.6.2 to guarantee reliable transmissions under variable channel condition for preallocated resources during the talk-spurt. The previous analysis indicates that the time and frequency aggressiveness can provide the BLER target only if the UE 's perceived SINR is reduced at least to 65% of the orginal value in downlink and 60% in uplink, in order to choose a more conservative MCS.

In the following simulations, time and frequency aggressiveness are set to corresponding value of 0.7 and 0.6 respectively. All the CQI and SRS strategies described in 3.5 will be considered in order to verify if and how the CQI and SRS reporting periodicity can influence the final BLER, the packets delays and the power consumption.

4.2.1 Single user scenario

The aim of this scenario is to analyze the SPS performance in a ideal case where only a single user has to be scheduled in the cell. That means the scheduler can use all the network resources to handle the user.

More frequent CQI reporting is supposed to assure better channel condition knowledge. With SPS scheduling the link adaptation is handled only at the beginning of each talk-spurt, so a too much frequent periodic CQI reporting should not be necessary. For this reason the period is set to 60 ms. In single user scenario, a CQI reporting at talk-spurt beginning, instead, can improve of around 20% the probability to achieve the BLEP target compared to the only periodic CQI reporting, as shown in figure 4.1. The same result can be reached with an aperiodic CQI reporting for the MCS reconfiguration. Hence a channel condition feedback at the beginning or during the talk-spurt allows a more reliable transmission exploiting the slow channel variations in downlink (figure 3.7 in section 3.4) and a conservative MCS selection that assures adeguate

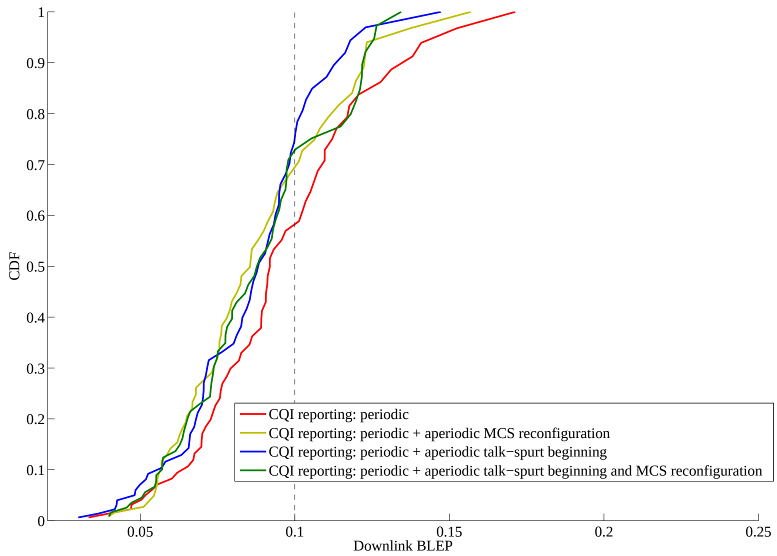


Figure 4.1: SPS - Downlink BLEP CDF, single user scenario

transmissions for a longer period than the one provided by the only periodic CQI reporting. In case of aperiodic CQI reporting both for talk-spurt beginning and MCS reconfiguration the final BLEP is similar to the previous cases but there are higher packets delays due to the needed time to require, receive and process the aperiodic CQI reporting.

It is important to underline that the proposed results have been obtained by a large number of simulations in order to realize a Monte Carlo statistical analysis, as explained in section 1.2.1. Due to the presence of high number of random events among the simulations, like network jitter, channel state variations, different CQI reporting measurement instants, different call traces, the described CQI reporting strategies can obviously show different behaviour case by case. Despite this, the results proposed can describe the average BLEP trend among the CQI reporting strategies.

As concern the uplink, the SRS reporting period is set to 20ms in order to follow the high variability of channel conditions. Figure 4.2 shows that the MCS reconfiguration with aperiodic SRS does not guarantee a high gain in BLEP target tracking. Furthermore this strategy could introduce more transmission delays due to the needed time to require and receive the SRS reporting.

4.2.2 Multiple users scenario

The purpose of the following analysis is to verify if an higher number of users in the cell can modify the SPS performance compared to the single user scenario. In this scenario we will consider 250 users at cell edge.

Comparing the BLEP CDF of figure 4.3 with the one in figure 4.1, for the downlink, and figure 4.2 and 4.4 for the uplink, there are not significant

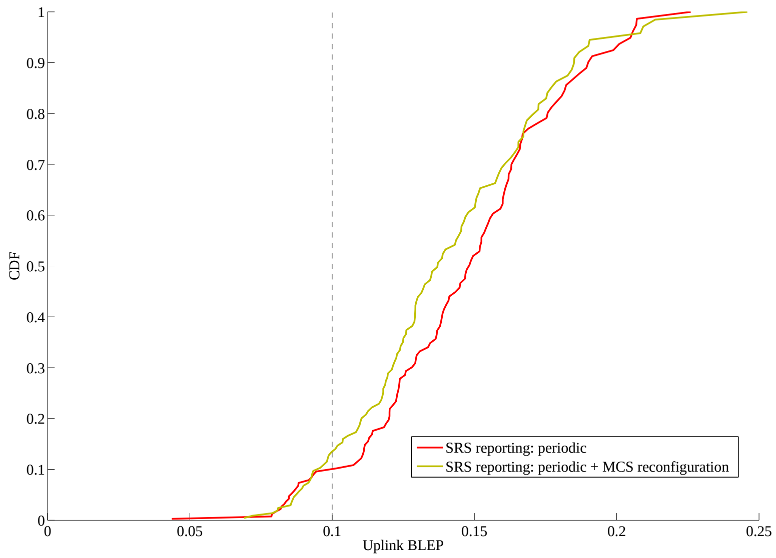


Figure 4.2: SPS - Uplink BLEP CDF, single user scenario

differences. Also in this case, the BLEP target is almost ensured in downlink but not in uplink.

Because of the small improvements provided in the previous analysis by the aperiodic SRS reporting in BLEP target tracking, compared to the periodic SRS reporting performance, we will leave out this case in the following DRX/DTX simulations.

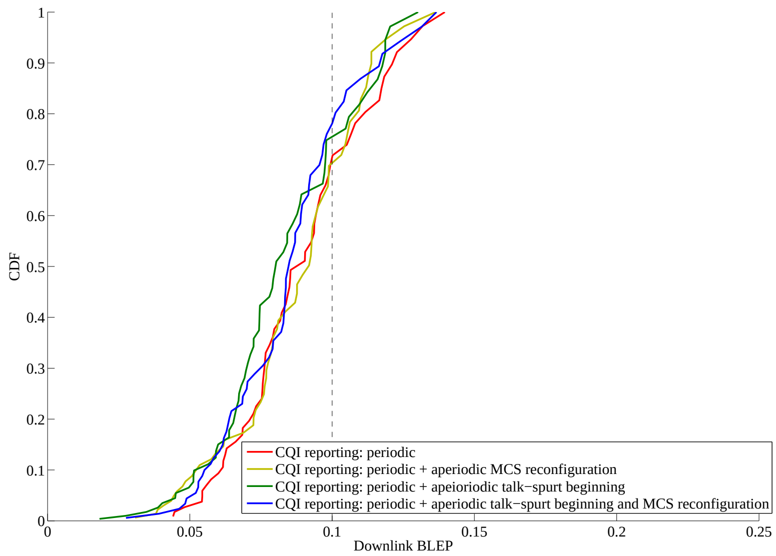


Figure 4.3: SPS - Downlink BLEP CDF, multiple users scenario (250 users at cell edge)

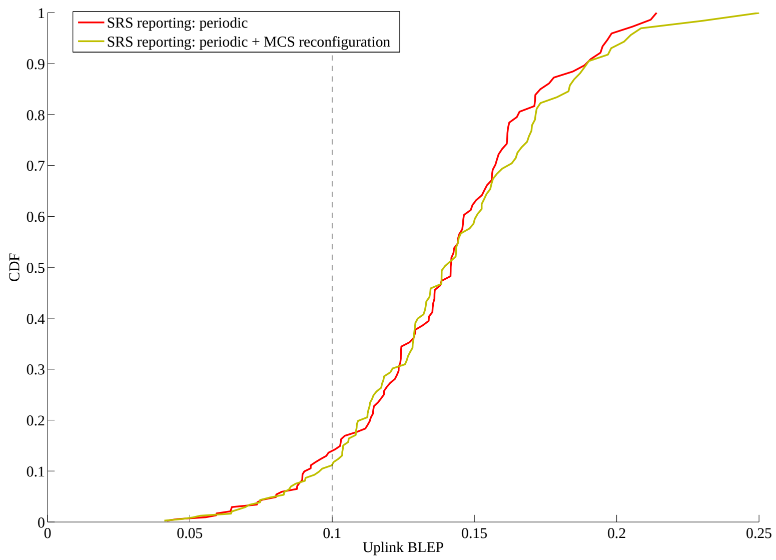


Figure 4.4: SPS - Uplink BLEP CDF, multiple users scenario (250 users at cell edge)

4.3 DRX/DTX impact

After a brief SPS performance introduction without DRX/DTX, the interaction between SPS and DRX/DTX will be evaluated. The purpose is to underline the effect of DRX/DTX on SPS performance, in terms of packets delays, and so user satisfaction, as well as the performance of SPS on power consumption, and how much the DRX/DTX can guarantee power saving.

The simulation and DRX/DTX parameters are shown in table 4.1

Table 4.1: Simulation parameters Sim #4.1

Simulation parameter	Value
Simulation Time	80 s
Loaded Users	1; 250
CQI reporting scheme	Wideband CQI
Bandwidth	10 MHz
BLER target	10%
Time Aggressiveness	$[0, 7]$
Frequency Aggressiveness	$[0, 6]$
CQI Reporting	60ms
SRS Reporting	20ms
DRX Long Cycle	20 ms
On-duration	1; 5; 7 ms

The long DRX cycle is set to 20 ms according to the AMR codec voice frame generation, the on-duration to 1, 5 and 7 ms. The value of 1 ms could be plausible in SPS because each packet is scheduled every 20 ms starting from the talk-spurt beginning without control signaling. Hence the UE can go to the active state in a precise instant of time only to receive or send the packet and check the PDCCH, without signaling activity.

4.3.1 Single user scenario

In section 4.2 it has been shown the gain introduced by aperiodic CQI reporting in BLEP target tracking for downlink transmissions. The benefits of BLEP reduction are lower network load and the possibility to reduce the time in active state in case of downlink retransmissions. Nevertheless there are not improvements on power saving (figure 4.5) due to the little impact that the downlink transmissions have on the final power consumption.

In fact, if we suppose that the UE receives 200 packets in downlink and sends 200 packets in uplink in 1000 TTIs, with 100 retransmissions both for uplink and downlink, the average power consumption, due to the transmitted and received packets, is 1.05 W. If retransmissions are reduced to 30 in downlink, the average power consumption is 0.8 W and the actual saving is only of 10 mW. This behaviour is associated to the higher uplink transmission power consumption (section 3.8). So, if we suppose to reduce also the uplink retransmissions, from 100 to 30, the final average power consumption is 0.8050 W. In this case the battery saving is almost 200 mW.

In this scenario, the DRX/DTX does not introduce high delays because all

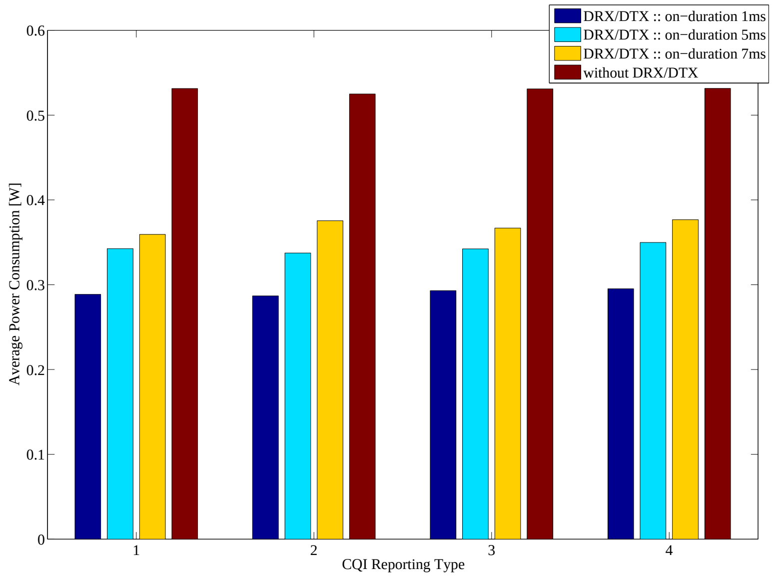


Figure 4.5: SPS - Average power consumption, single user scenario

time/frequency resources are available. The packets dropping events (figure 4.6) could be due to the needed time to deliver them that includes DRX/DTX plus retransmissions delays.

4.3.2 Multiple user scenario

In the following analysis the effect of multiple users scenario will be shown to verify if there is an impact to power consumption and user satisfaction when the scheduler has to handle 250 users at the cell edge.

Comparing the power consumption in figure 4.7 with the one in figure 4.5 it is possible to verify that the average values are almost the same in both cases. So, 250 users in the cell can not influence the DRX/DTX behavior and compromise the SPS performance in downlink, as it can be seen from the low number of dropped packets in figure 4.8.

As concern the uplink, the delays (figure 4.9) are higher than the downlink ones because, as explained in section 3.6.2, the uplink requires more resources to ensure a reliable transmission and, with 250 users, the number of free PRBs is significantly reduced. This means that, because the DRX/DTX activity reduces the time instants where the UE can be scheduled, the resources can be already allocated to other users when the UE goes to active state.

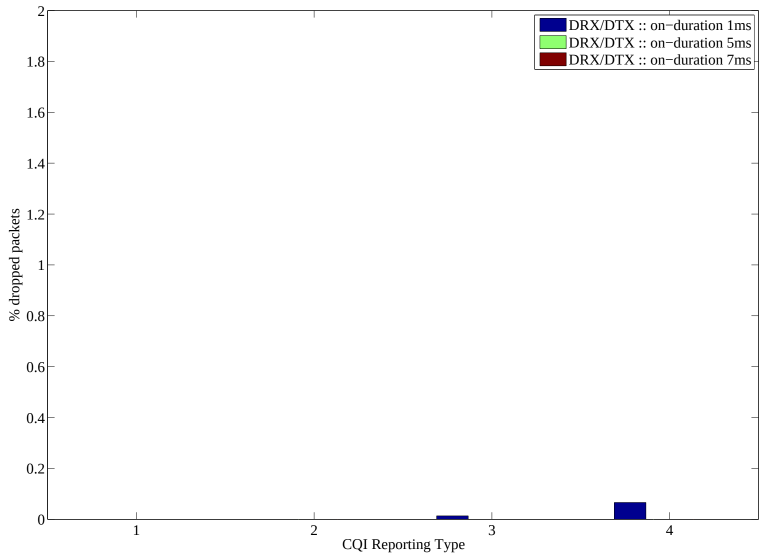


Figure 4.6: SPS - Downlink packet drop probability, single user scenario

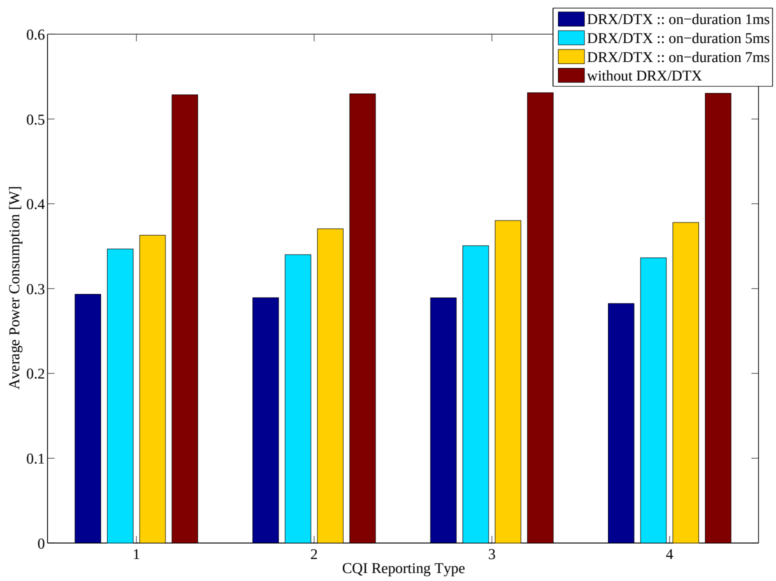


Figure 4.7: SPS - Average power consumption, multiple users scenario (250 users at cell edge)

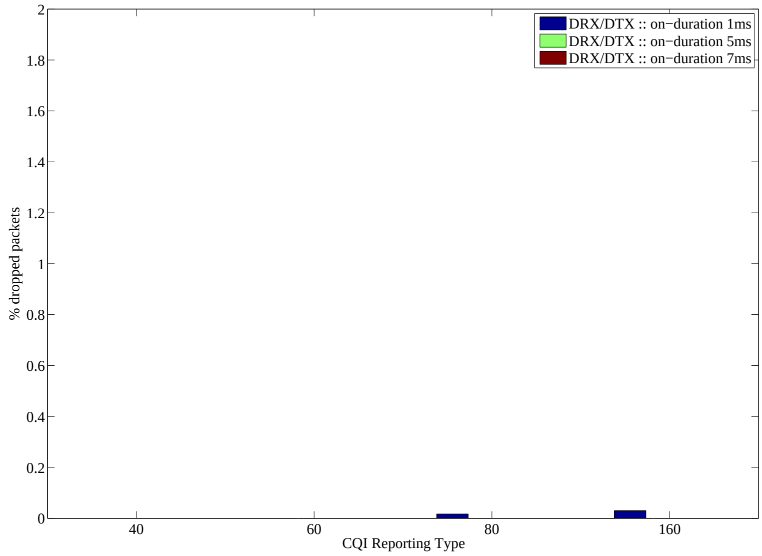


Figure 4.8: SPS - Downlink packet drop probability, multiple users scenario (250 users at cell edge)

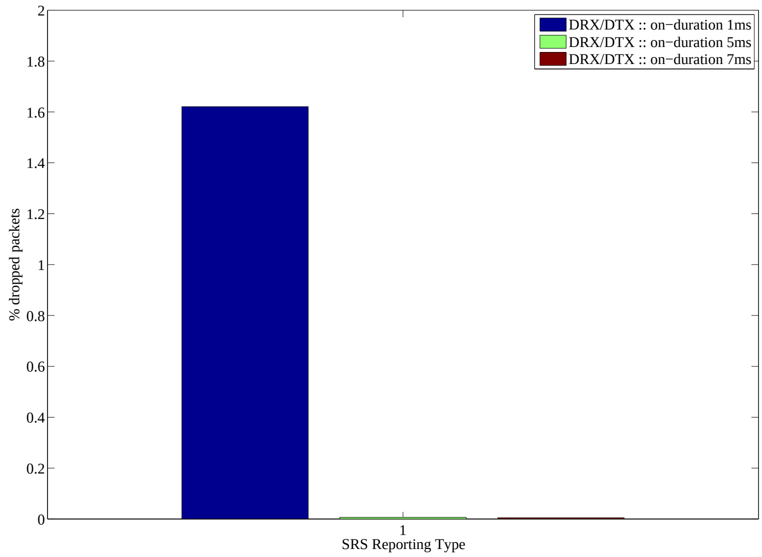


Figure 4.9: SPS - Uplink packet drop probability, multiple user scenario (250 users at cell edge)

4.4 Short DRX/DTX

The use of the short DRX/DTX, compared to the static DRX/DTX, is another interesting DRX/DTX functionality to analyze. The short DRX/DTX purpose is to further reduce the time instant in active state when the UE is not receiving/sending data or control signaling for long periods.

Also in this case, the simulations will be carried out both for single and multiple scenario. The parameters are shown in tab. 4.2

Table 4.2: Simulation parameters Sim #4.2

Simulation parameter	Value
Simulation Time	80 s
Loaded Users	1; 250
CQI reporting scheme	Wideband CQI
Bandwidth	10 MHz
BLER target	10%
Time Aggressiveness	$[0, 7]$
Frequency Aggressiveness	$[0, 6]$
CQI Reporting	60ms
SRS Reporting	20ms
DRX Long Cycle	40; 60; 80; 160 ms
Short DRX Cycle	20 ms
On-duration	1; 5; 7 ms

All the CQI reporting strategies have been considered. However, only the strategies 1 and 3 of those defined in section 3.5 will be shown. The aim of that choice is to represent the DRX/DTX behaviour for the “best” and “worst” CQI reporting strategies, as shown in section 4.2.

4.4.1 Single user scenario

The following results will be divided in two set of analysis: CQI reporting types 1 and 3, varying in each case the long DRX cycle and the on-duration period. In general, the short DRX/DTX seems to not provide significant power saving in this scenario (figures 4.10 and 4.11) if compared to the power consumption values of the static DRX/DTX case (figure 4.5). Further, the same behavior can be observed comparing the power consumption with long DRX cycle of 40 ms to the one with long DRX cycle of 160 ms: there is only a slight difference between them.

The reason of these results can be found if we consider the behaviour of two users engaged in bidirectional conversation during a generic call. Since we are considering both downlink and uplink, the UE terminal is most of the time in active state with data reception or transmission. In fact, even if the user is quite, the second user, is probably talking. In this way the UE terminal will receive voice packets every 20 ms and send SIDs every 160 ms, according to the AMR codec frame generation. Hence the short DRX cycle can be switched off only during the mutual silence events but, statistically, there are few mutual silence events in a bidirectional conversation (see section 3.2).

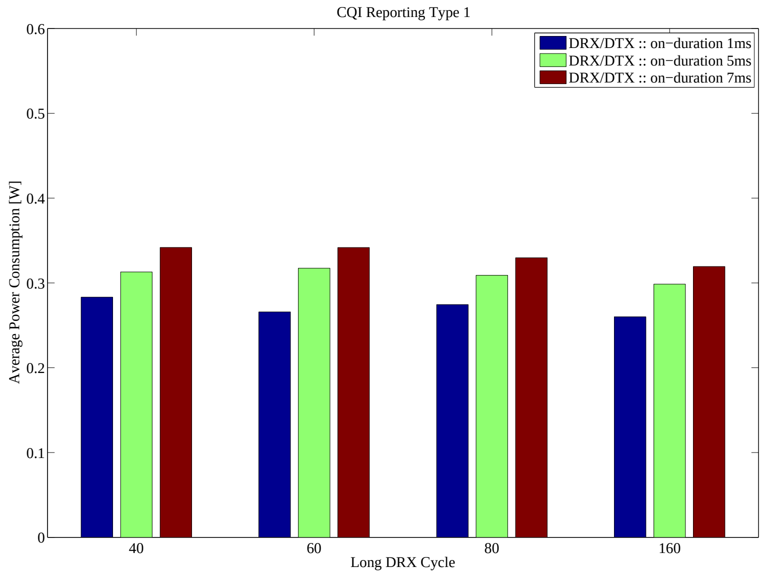


Figure 4.10: SPS - Average power consumption, single user scenario, CQI reporting type 1

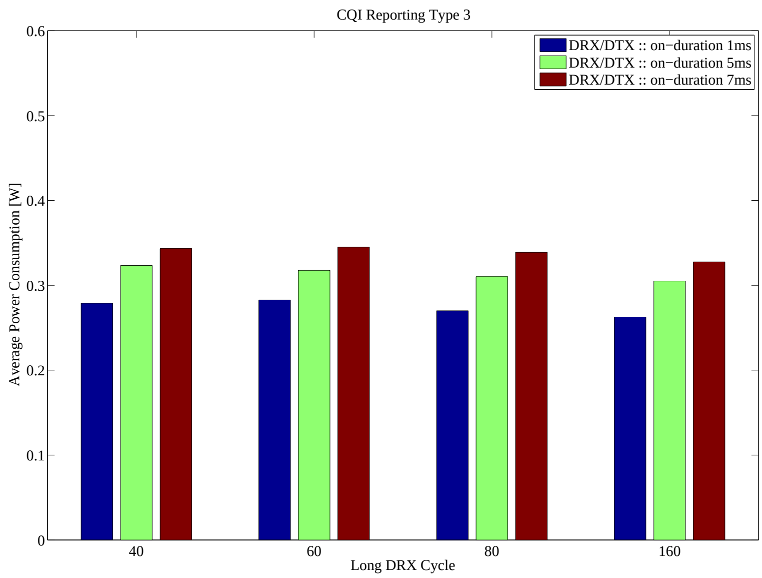


Figure 4.11: SPS - Average power consumption, single user scenario, CQI reporting type 3

The only short DRX remarkable effect consists in higher delays and thus more dropped packets (figures 4.12 and 4.13). That behaviour is more explicit in downlink than in uplink. In fact, in downlink the UE can be scheduled only when it goes back to active state. This means that, if the first packet of talk-spurt reaches the buffer few TTIs after the end of the last on-duration, the UE could be scheduled only after around 120 ms. In this period at least the first 4 packets will be dropped by the eNodeB. Furthermore these delays can be worsened if an aperiodic CQI reporting is required at the talk spurt beginning, as it is possible to see comparing the figures 4.12 and 4.13.

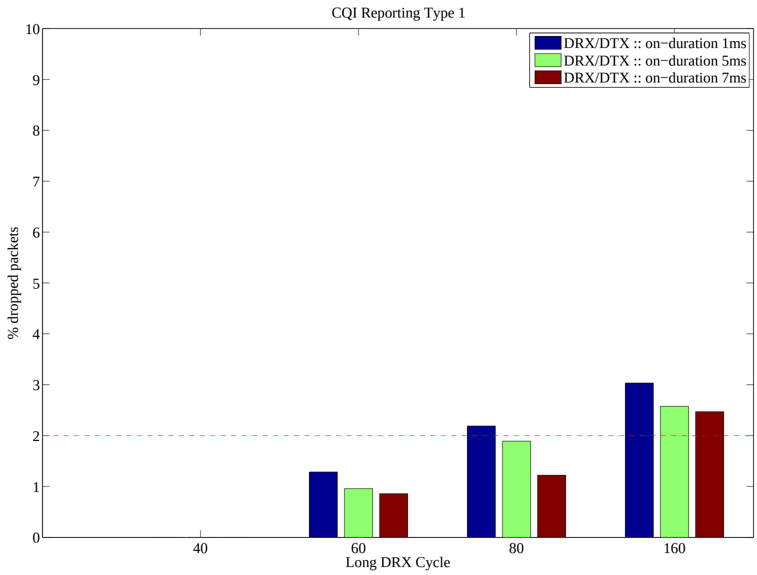


Figure 4.12: SPS - Downlink packet drop probability, single user scenario, CQI reporting type 1

For the uplink, instead, the UE will be scheduled after few TTI the scheduling grant. In single user scenario, since all resources are free, there are not packets delays.

4.4.2 Multiple users scenario

The multiple users scenario shows a slight increase on power consumption (figures 4.14 and 4.15) if compared to the single user scenario (figures 4.10 and 4.10), both for CQI reporting type 1 and 3. Also in this case an extended period of long DRX cycle can not assure a clear power consumption reduction for the same reason explained before.

As concern the downlink delays, they are further increased than the single scenario, because of the presence of more users in the cell. This means that the resources can be completely allocated to other users when the UE goes to active state. So if the on-duration is 1 ms the scheduler can not serve the UE in the current and following TTIs, as can happen for longer on-duration (5 ms and 7

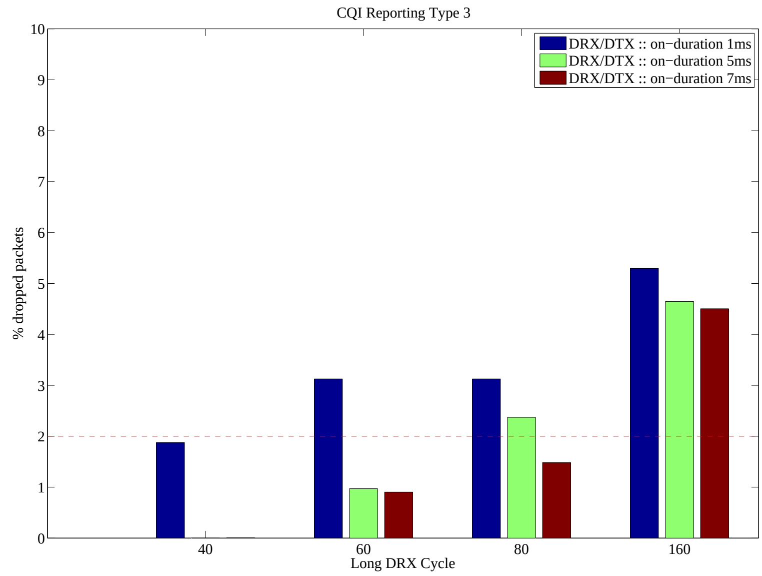


Figure 4.13: SPS - Downlink packet drop probability, single user scenario, CQI reporting type 3

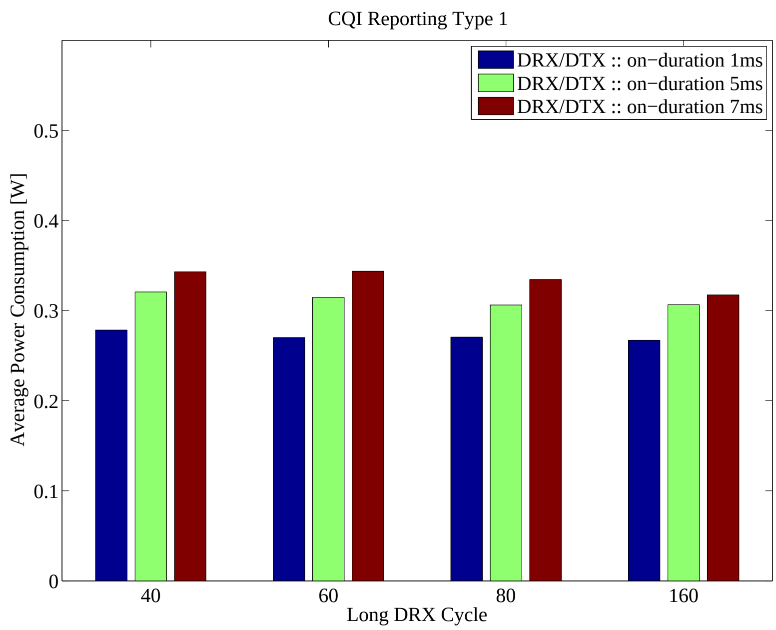


Figure 4.14: SPS - Average power consumption, multiple users scenario (250 users at cell edge), CQI reporting type 1

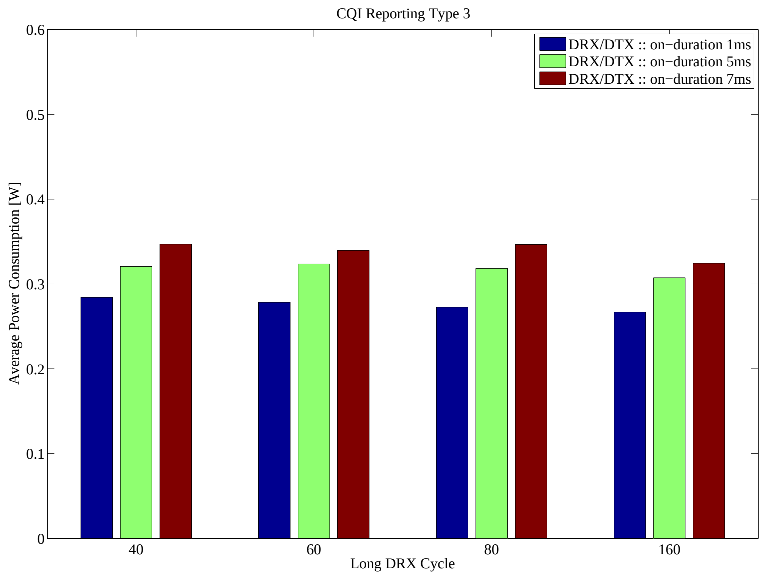


Figure 4.15: SPS - Average power consumption, multiple users scenario (250 users at cell edge), CQI reporting type 3

ms), but only in the next active state (4.16). This behaviour can introduce high packets delays.

Moreover, these delays can increase if the long DRX is set to 160 ms, for the same reason explained in 4.4.1, and in this case the number of dropped packets can increase over the user satisfaction threshold of 2% (figures 4.17 and 4.18).

Also in this scenario, the uplink delays are low: the UE can be scheduled after only few TTIs after the scheduling grant (figure 4.19), exploiting the chance to anticipate the DTX active state with the short DRX/DTX.

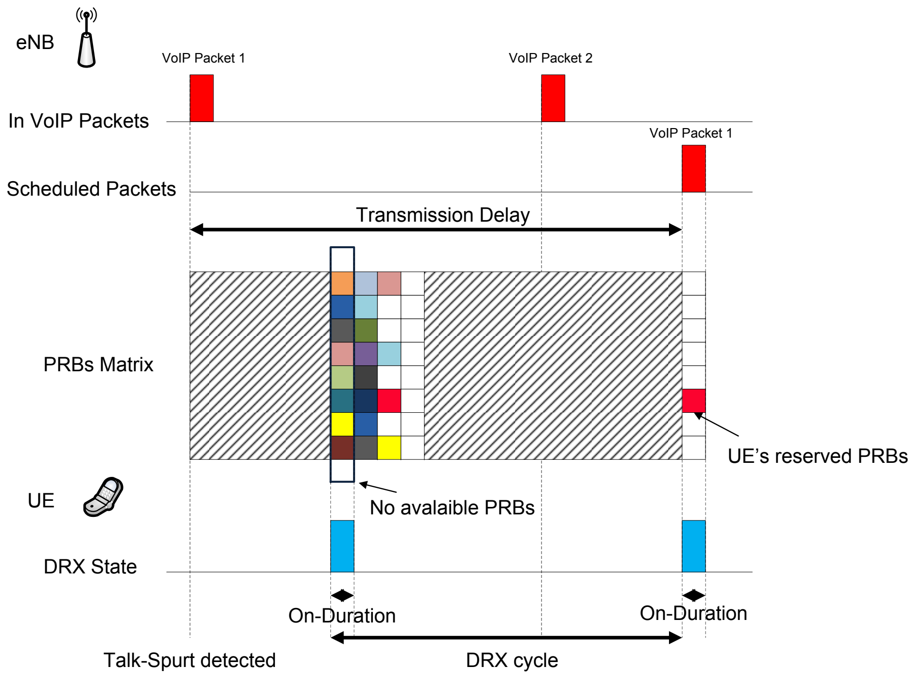


Figure 4.16: Downling UE scheduling with static DRX/DTX

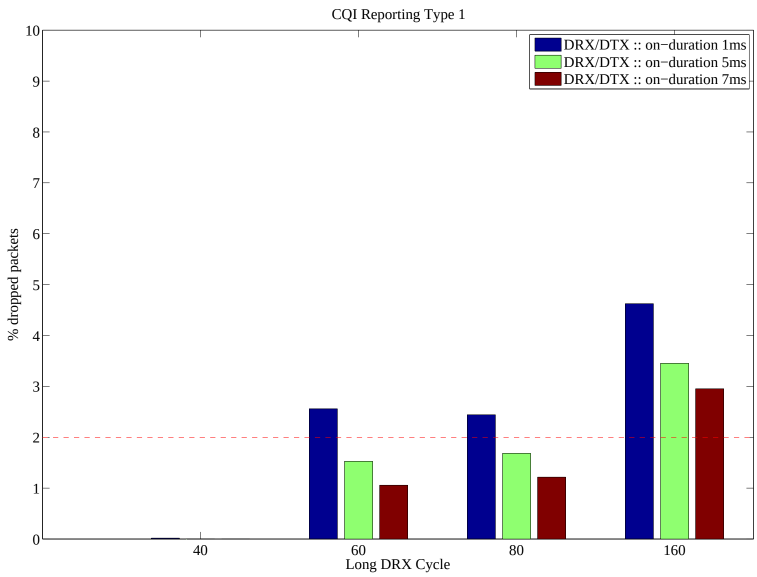


Figure 4.17: SPS - Downlink packet drop probability, multiple users scenario (250 users at cell edge), CQI reporting type 1

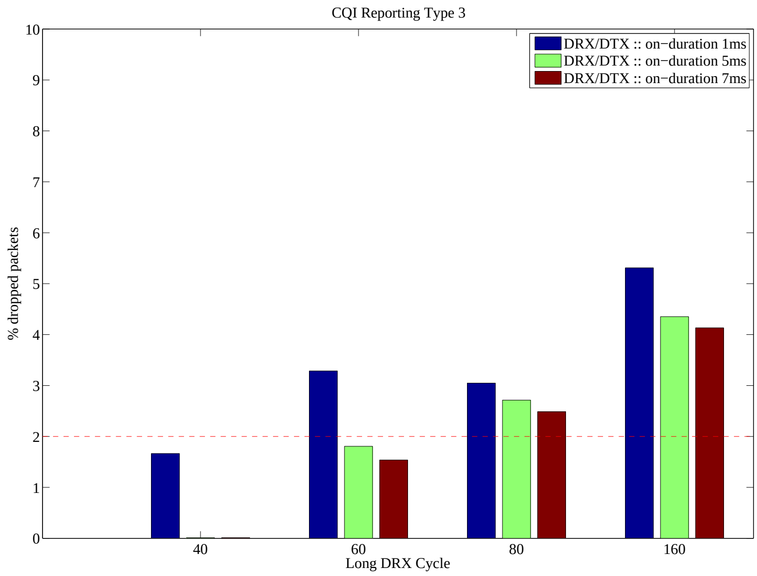


Figure 4.18: SPS - Downlink packet drop probability, multiple users scenario (250 users at cell edge), CQI reporting type 3

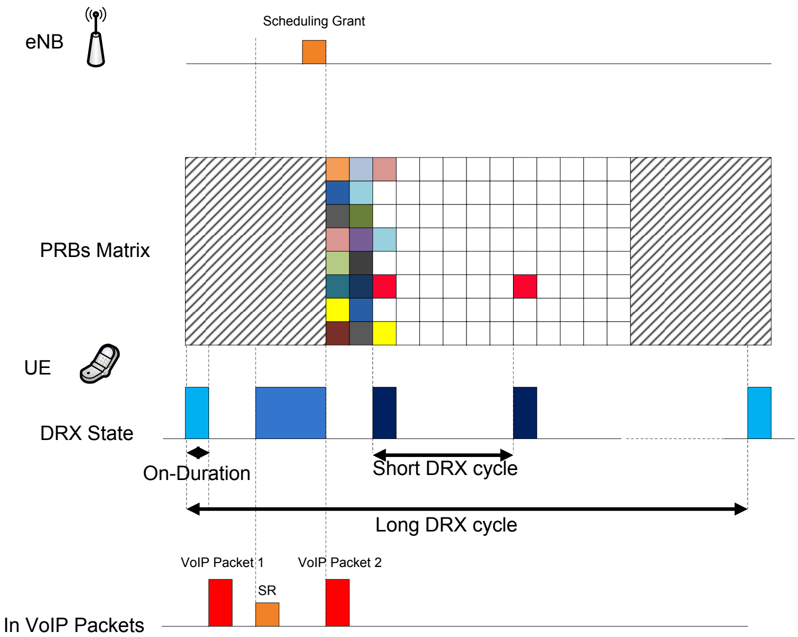


Figure 4.19: Uplink UE scheduling with short DRX/DTX

4.5 Main outcome

The shown results confirm the power saving gain provided by the DRX/DTX functionalities. In fact, there is a difference of around 200 mW/TTI between power consumption without DRX/DTX and with static DRX/DTX. The effect of longer on-duration period can obviously change the final power consumption, but at the same time can reduce packets scheduling delays in case of higher users in the cell, both for downlink and uplink.

The short DRX/DTX utility can not assure a further power saving because of the absence of long mutual silence periods between two talk-spurts. This leads the UE to a continuous data reception/transmission during the whole call. Moreover, the use of extended long DRX cycle decreases the downlink SPS performance in terms of number of dropped packets. The uplink, instead, seems to guarantee good performance from this point of view. When the UE has to transmit, it can wake up without waiting a new on-duration event and can be scheduled in the first available TTI, triggering the beginning of the short DRX/DTX in that instant.

Chapter 5

DRX/DTX with DS

In the following section we will analyze the behavior of the previously described dynamic packet scheduler in several system level simulations. DS is the most commonly used scheduler for traffic that shows bursty dynamic, and non repetitive patterns. The characterization of VoIP traffic suggests that semi-persistent packet scheduling should fit better the periodicity and fixed size of VoIP packets, but we want to analyze if better tracking of channel conditions achievable with DS can show fine effects over the number of retransmissions and therefore on power consumption.

We will analyze the general behavior of DS in single and multiple users scenarios, where other 250 users, similarly located at the cell edge, require the same type of service. The simulations will involve both static DRX/DTX and short DRX/DTX, with different settings for channel quality feedback reporting period.

The chosen performance evaluation metrics are packets delivery delays and average power consumption.

5.1 Modeling Assumptions

Like SPS case, the considered UE is moving along the cell edge at the speed of 3km/h, with radio channel conditions previously explained in section 3.4 based on traces that assume full interference. The simulation is performed on a single user basis, but other users and thus network load, are emulated filling the time/frequency grid with data from previous asynchronous simulations. In this way results for single user and 250 users scenarios are also presented. The single user scenario considers a completely empty time/frequency allocation table, so the time domain scheduler has not any problem allocating each new packet in the first possible TTI. The frequency domain scheduling exploits an optimal situation too. Since we assume a full CQI reporting, the scheduler can assign the best PRBs for each transmission.

In the following simulations we will not consider TTI packets bundling because not implemented.

The core functionality of dynamic packet scheduler, that differentiates it from SPS, is the continuous adaptation to channel conditions, in terms of number of assigned PRBs and their position among the whole channel. While this solution assures better BLEP target tracking at the level of single transmission (while SPS can assure it as average on the whole talk-spurt), it generates much more quality reports and signaling in general. In fact, the scheduler core is run every TTI, without saving an allocation state: every new packet detection triggers a new scheduling decision that requires frequent activity of channel measurement entities.

Section 3.6.1 accurately describes the functionalities of dynamic scheduler.

5.2 Dynamic scheduling performances without DRX/DTX

The BLEP target for MCS decisions is fixed to 10% for both downlink and uplink. Considering the iterative nature of dynamic scheduler, it is possible to define a main differece in the importance of the signaling activity performed inside and outside talk-spurt periods. However the amount of this signaling depends on configuration, SPS bases its scheduling *a priori* choice on the informations received during non-talks-purt periods, and does not take advantages of the ones reported during talk-spurt ones. On the contrary, DS scheduler exploits the whole talk-spurt duration to continuosly adapt each packet transmission to the channel quality trend, basing its choices on the quality reporting activity. It can be said that the SPS's control action is performed *offline* with the talk spurt and the DS's one is *online* instead.

The first analysis executed to evaluate the performances of dynamic scheduling with VoIP traffic does not include any type of power saving policy, so the RF modem is kept on for the whole duration of the simulation. This means that the UE responsiveness is immediate, there is not any type of delay due to sleeping states of the UE.

This kind of analysis is performed to simply evaluate how much DS channel quality conditions tracking is valuable, to verify if it is worth to be used despite its overhead.

5.2.1 Single user scenario

As said before, the single user scenario considers an ideal case where the UE can be scheduled in a short amount of time, and over the best frequencies where it is not affected by signal degradation phenomena.

Table 5.1 summarizes the simulation parameters used to evaluate DS performances in downlink.

OLLA A_{UP} offset is set to 0.5 dB and we choose two CQI reporting periods, 10 and 20 ms, that show good tracking properties (figure 3.11). This choice is based on the OLLA behavior analysis described in section 3.6.1 and some DRX/DTX considerations that will be explained in the next sections. CQI selectivity is 3 PRBs, the maximum value specified by 3GPP.

Both of the simulations results show (figure 5.1) a very fine tracking of the selected BLEP target. This means that the OLLA control is correctly

Table 5.1: Simulation parameters Sim #5.1

Simulation parameter	Value
Simulation Time	80 s
Users in the system	1
CQI selectivity	Full CQI
CQI granularity	3 PRBs
Bandwidth	10 MHz
BLEP target	10%
A_{up}	0.5dB
CQI Reporting	$[10, 20]$ms

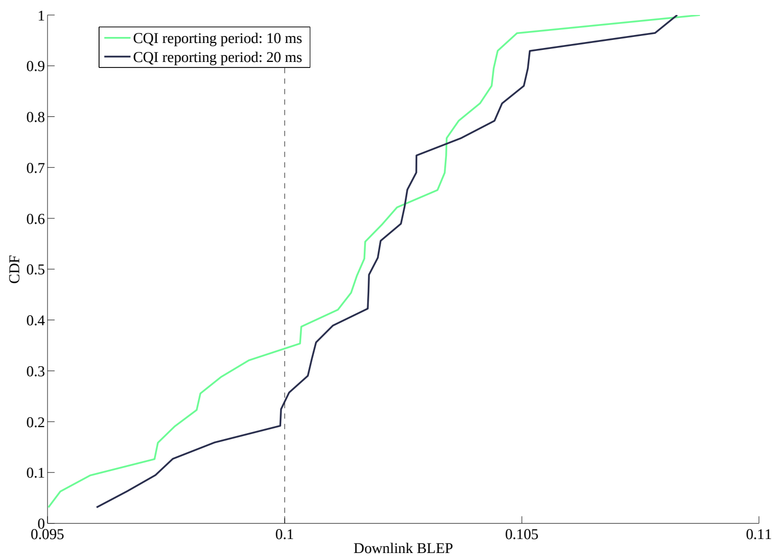


Figure 5.1: DS - Downlink BLEP CDF, single user scenario

parameterized: averaging the results of simulations, the 100% of values are included in the target ±1%. It can be seen that the curve with 10ms of CQI reporting period shows a better behavior than the other: this is due to the benefits that OLLA receives with a more frequently refreshed memory of channel quality indicators. However these improvements have to be evaluated to verify if they are worth, in fact they consume signaling resources and require more sleep-to-active transitions if DRX-DTX is enabled.

DS uplink evaluation is conducted with the same parameters of downlink (table 5.2), tuning the SRS reporting period to 10 and 20 ms for the same reasons. OLLA uplink control shows an average good behavior with these parameters as represented in figure 3.12.

Table 5.2: Simulation parameters Sim #5.2

Simulation parameter	Value
Simulation Time	80 s
Users in the system	1
SRS selectivity	Full CQI
SRS granularity	3 PRBs
Bandwidth	10 MHz
BLEP target	10%
A_{up}	0.5dB
SRS Reporting	$[10, 20]$ms

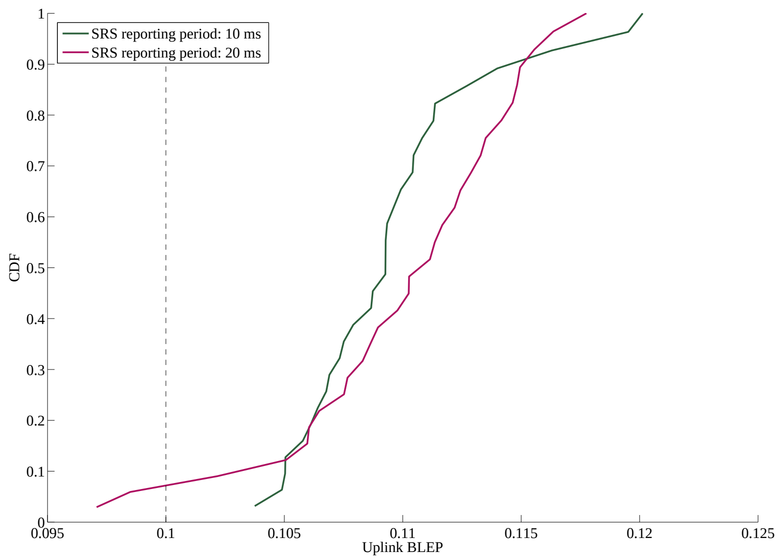


Figure 5.2: DS - Uplink BLEP CDF, single user scenario

Figure 5.2 shows slight worse results mainly due to higher values of variance of uplink channel quality that make the accurate tracking of it more difficult. However, OLLA control is still capable to maintain the overall BLEP of transmissions around the preferred target. Also in this case a more frequent channel quality reporting improves the tracking, in fact the 10ms CDF if steeper than the other.

5.2.2 Multiple users scenario

Next analysis involves a system with a large number of concurrent users. The purpose is to evaluate how much is the BLEP tracking decreased when the scheduler meets difficulties finding the needed empty PRBs in the time/frequency grid. Since we assume packet segmentation in different TTIs is disabled, the time domain scheduler must find a TTI row with enough free PRBs to transmit the whole packet with the chosen MCS. The "always on" particular UE state of these simulations, however, makes space to a simpler choice for the scheduler because of the large set of available TTIs in which it can search for available resources.

Table 5.3: Simulation parameters Sim #5.3

Simulation parameter	Value
Simulation Time	80 s
Users in the system	250
CQI selectivity	Full CQI
CQI granularity	3 PRBs
Bandwidth	10 MHz
BLEP target	10%
A_{up}	0.5dB
CQI Reporting	$[10, 20]$ms

The network load given by other 250 users allocations seems not to represent an issue in downlink and the very low values of BLEP imply the lack of significant delays, so that the considered user is always satisfied. The slow variance of downlink channel trace allows the channel reporting entity to correctly estimate the conditions, giving results that can be proficiently used for a pretty high number of TTIs: the scheduling delay, induced by already occupied resources, does not influence the BLEP tracking a lot. Moreover good tracking means a correct estimation of needed PRBs, without waste of them. This result in a time/frequency grid that, however filled with 250 users, has still enough free space to not significantly decrease the scheduler performance.

The results are quite different in uplink. The highly variable uplink trace causes more retransmissions than downlink and an overall poorer BLEP tracking. Since each transmission is subject to a general higher number of retransmission, each of which has a growing number of needed PRBs, causes a premature filling of the allocation grid. In every simulation we have observed that the uplink high load percentage, i.e. inability to accept other users, leads the system to outage events.

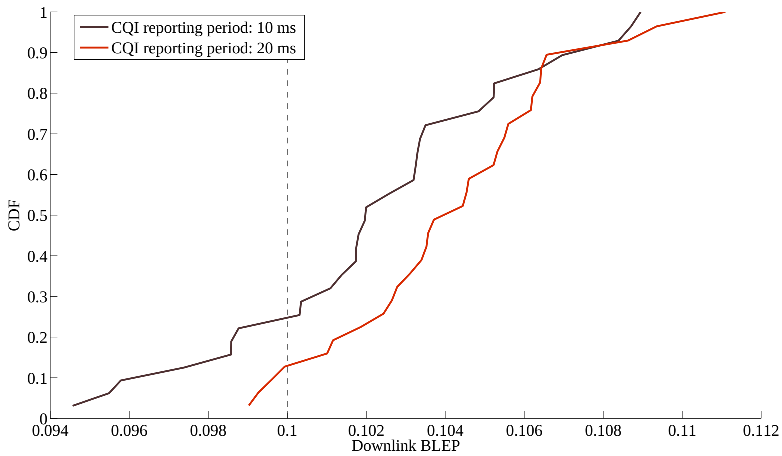


Figure 5.3: DS - Downlink BLEP CDF, multiple users scenario (250 users at cell edge)

Table 5.4: Simulation parameters Sim #5.4

Simulation parameter	Value
Simulation Time	80 s
Users in the system	250
SRS selectivity	Full CQI
SRS granularity	3 PRBs
Bandwidth	10 MHz
BLEP target	10%
A_{up}	0.5dB
SRS Reporting	$[10, 20]$ms

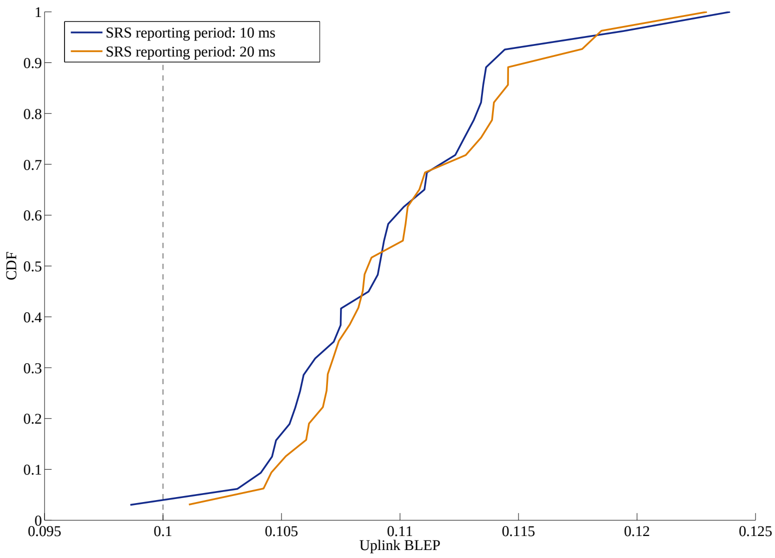


Figure 5.4: DS - Uplink BLEP CDF, multiple users scenario (250 users at cell edge)

5.3 DRX/DTX impact

The following simulations show the impact of DRX/DTX functionality on UE's power consumption. The tuning of DRX/DTX is responsible of several variations of these values, so we analyze which is the best adoptable solution. It's important to evaluate the further gains that short DRX/DTX activation can provide, and if it is a valuable technique to be used with VoIP traffic.

The following simulations set is performed with short DRX/DTX functionality disabled. This should produce a general improvement in power consumption, however independent from talk spurts call parts. The channel quality reporting, in downlink and uplink is forced to be synchronized with DRX/DTX cycle. The latter is set to 20 ms to allow the scheduler to smoothly flush the packet queue, since packets bundling is not enabled.

5.3.1 Single user scenario

The results produced with the introduction of DRX/DTX activity are compared with those of an *always-on* UE. The reason of the on-duration values choice is based on some scheduling considerations. Especially in downlink, a too small on-duration timer (less than 5 ms) would force the UE to go to sleep state before being able to receive the packet, counting the implicit scheduling delays. The $>$ 5 ms on-duration enables the UE to send the CQI report and receive the packet in a single DRX cycle activity, allowing the eNodeB to effectively schedule the transmission without waiting for another DRX cycle. Higher on-duration values should imply lower power consumption gains, penalizing the whole DRX/DTX functionality.

Table 5.5: Simulation parameters Sim #5.5

Simulation parameter	Value
Simulation Time	80 s
Users in the system	1
CQI/SRS selectivity	Full CQI
CQI/SRS Reporting	20ms
DRX/DTX Cycle	20ms
short DRX/DTX	disabled
on-duration	$[5, 7]$ms

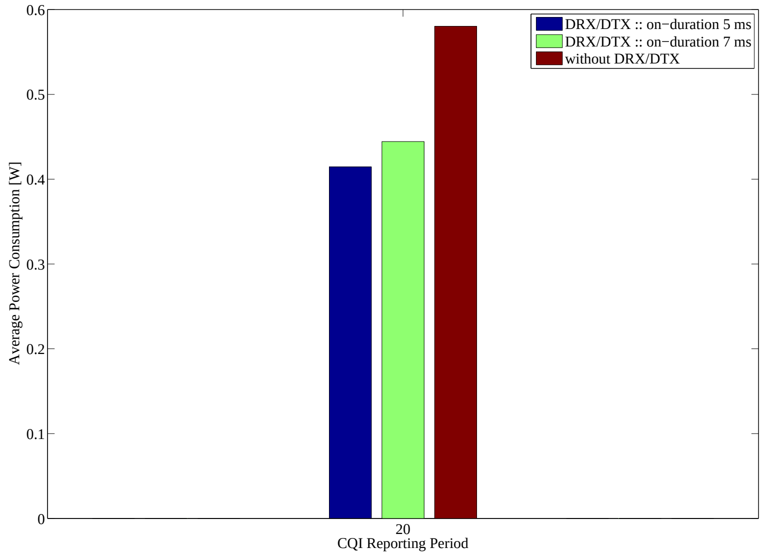


Figure 5.5: DS - Average power consumption, CQI/SRS periodicity 20ms, single user scenario

The single user scenario results confirm the average gains in power saving due to the sleeping pause of the UE. The difference of 2 TTIs in on-duration timer, however, gives a really slight power consumption difference, due only to the actual additional active TTIs (+66% of active without data TTIs).

5.3.2 Multiple users scenario

Table 5.6: Simulation parameters Sim #5.6

Simulation parameter	Value
Simulation Time	80 s
Users in the system	250
CQI/SRS selectivity	Full CQI
CQI/SRS Reporting	20ms
DRX/DTX Cycle	20ms
short DRX/DTX	disabled
on-duration	$[5, 7]$ms

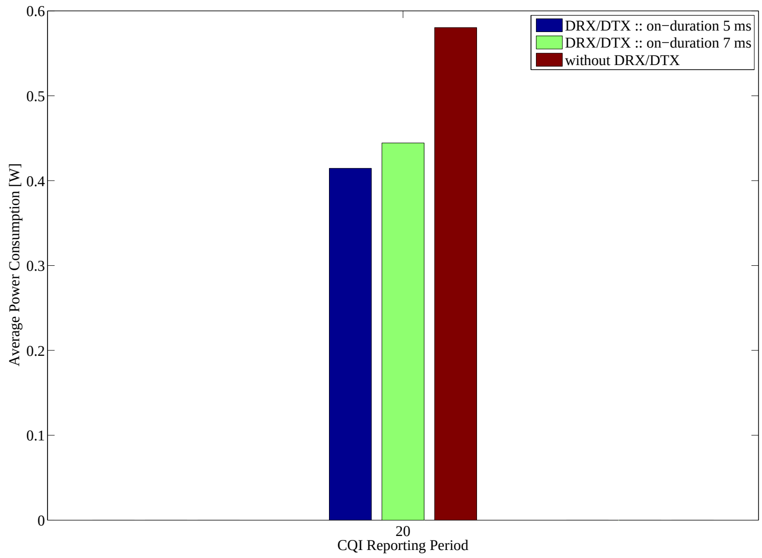


Figure 5.6: DS - Average power consumption, CQI/SRS periodicity 20ms, multiple users scenario (250 users at cell edge)

The more loaded scenario shows pretty the same results in power consumption savings compared to the disabled DRX/DTX case. However the difference between 5 and 7 ms on-duration timers is more evident: the reason of this behavior has to be found in the scheduling activity described in the previous paragraph. The presence of an high number of users in the systems reduces the chance to find free PRBs before the on-duration ending. This phenomenon is more accentuated when the set of eligible TTIs is very small, i.e. 5 TTIs. The direct consequence of this, is the postponement of the allocation to the next DRX/DTX cycle. If this referral happens more times, the eNodeB itself drops the packet because it is not able to guarantee a delay lower than 50ms. The lack of signaling and transmission of these packets represents a slight but significant reduction of consumed power.

This can be demonstrated in figure 5.7, that shows the significant increase of dropped packets percentage when on-duration timer is set to low values.

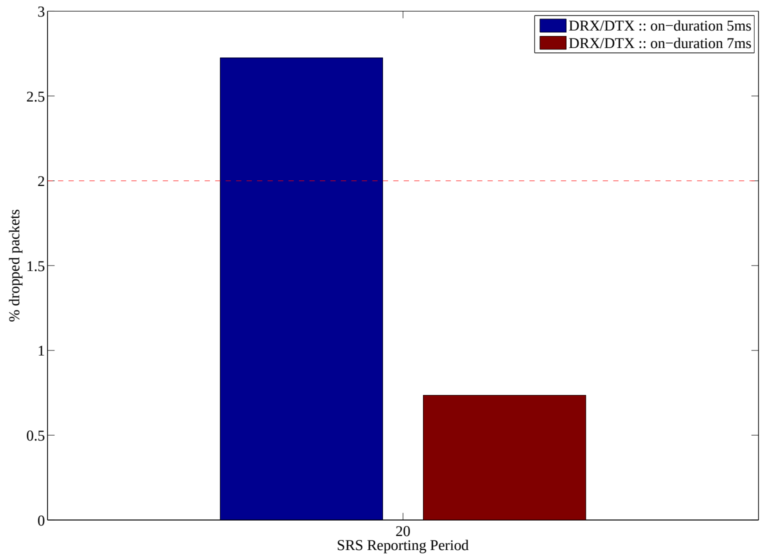


Figure 5.7: DS - Uplink packet drop probability, multiple users scenario (250 users at cell edge)

5.4 Short DRX/DTX

The introduction of short DRX/DTX functionality could optimally adapt the RF modem activity to the VoIP traffic characteristics. In fact, it would be obvious to set a long DRX/DTX cycle that fits the SID packets frequency and the short cycle to the AMR codec activity. These choices, however, are the best solutions only in ideal conditions, with a low loaded network and good channel conditions. In real scenario a lot of random factors, like jitters, retransmissions bursts, fully loaded TTIs, and so on, can influence and worsen the transmission process.

We will analyze the performances of the systems in terms of user satisfaction and power consumption. The satisfaction is evaluated measuring the number of packets dropped by the source or by the receiver because of a scheduling delay higher than 50 ms. The analysis are performed with a single or 250 users in the system, with several values of channel quality reporting periods.

5.4.1 Single user scenario, CQI reporting period 10ms

The first set of simulations involves a single user with short DRX/DTX enabled. The CQI reporting is set to 10 ms.

Table 5.7: Simulation parameters Sim #5.7

Simulation parameter	Value
Simulation Time	80 s
Users in the system	1
CQI/SRS selectivity	Full CQI
CQI/SRS Reporting	10ms
DRX/DTX Cycle	$[40, 60, 80, 160]$ms
short DRX/DTX cycle	10ms
on-duration	$[5, 7]$ms

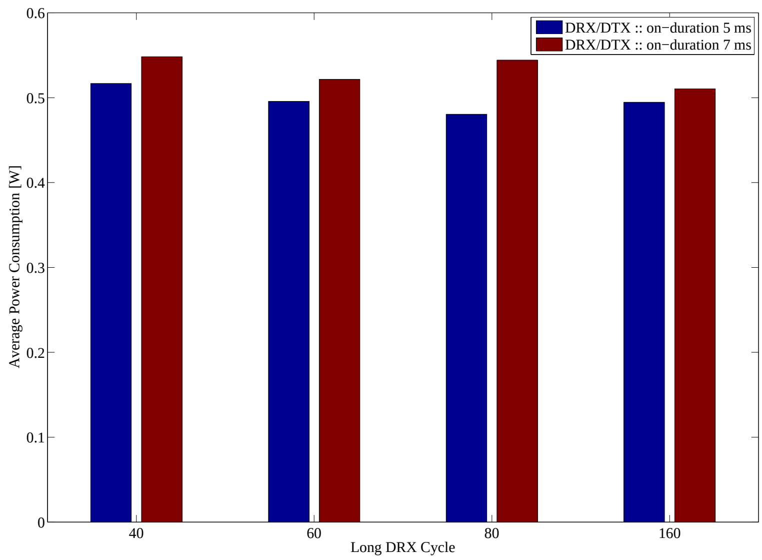


Figure 5.8: DS - Average power consumption, shortDRX, CQI/SRS periodicity 10ms, single user scenario

Figure 5.8 - The results are quite self-explaining: a longer DRX/DTX cycle results in very poor power consumption gains. The main reason of this behavior should be searched for in the particular traffic type we are analyzing. The long DRX/DTX condition is triggered only when no voice data have to be send both in downlink and uplink. This corresponds to our *mutual silence* speech states, and reviewing table 3.3 it is clear that this condition is reached few times and for small periods.

Figure 5.9 - The low relevant differences in power consumption given increasing the DRX/DTX long cycle to 160 ms, i.e the periodicity of SID packets, have

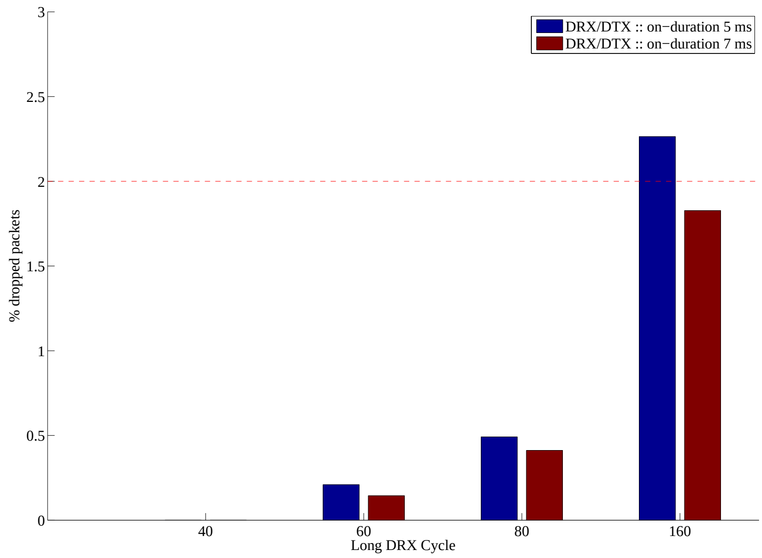


Figure 5.9: DS - Downlink packet drop probability, shortDRX, CQI/SRS periodicity 10ms, single user scenario

though quite a bit impact on downlink packet drop probability. The threshold of 2% is not reached, but considering that this is a single user scenario it can be predicted that the results obtained with so high long DRX/DTX cycles are inacceptable.

The reason behind these high packet drop probabilities resides in the general high time a packet have to spend in the transmission queue if a DRX/DTX cycle is just finished. The high probabilities are, in fact, the result of packet discarding acted by the eNodeB, that can be very frequent with 160 ms long cycle: considering that the eNodeB drops a packet if its scheduling delay is higher than 50ms, if the worts case of a just finished DRX/DTX cycle happens, there is a time frame of 110 ms (or even more if eNodeB takes into account scheduling delays too) that corresponds to a potential amount of 5 dropped packets. Table 3.4 indicates that a talk-spurt lasts about 1.8 seconds, i.e. 90 packets: this means a loss of 5,6% of talk spurt's data. The effect of a 160 ms long DRX/DTX cycle is a truncation of each downlink talk spurt beginning.

Figure 5.10 - The uplink transmission is not affected by this effect. This is due to the different transmission process actuated when new data are available at the UE. In fact, the UE itself can "break" the long sleep state triggering the RF modem activation, without waiting for a planned activation. This asynchronous process does not cause any kind of delay at the source of the transmission. A very low packet drop probability is detected with 160ms long DRX/DTX cycle, and this should be imputable to first packet unsuccessful of talk spurt transmissions due to initial bad channel estimations. The OLLA control, in fact, is rarely run, so it is not able to fix a good starting estimation at the beginning of talk spurt.

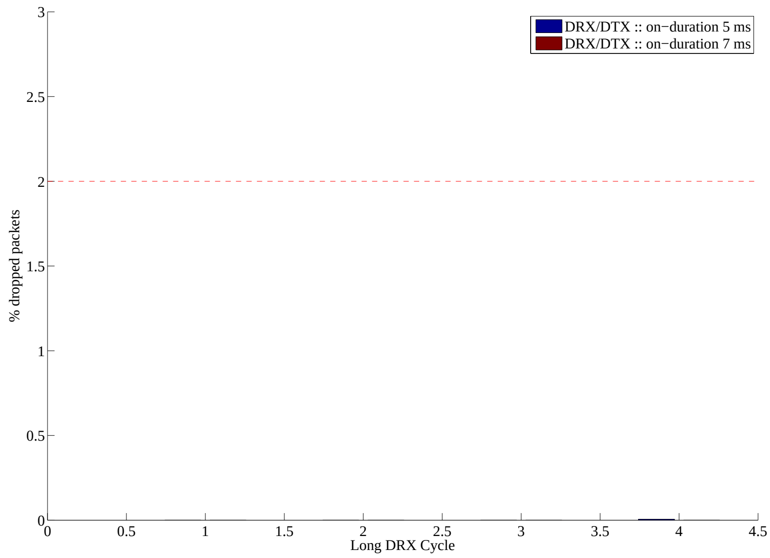


Figure 5.10: DS - Uplink packet drop probability, shortDRX, single user scenario

5.4.2 Multiple users scenario, CQI reporting period 10ms

The multiple users scenario emulates an high traffic load on the network that force the scheduler to handle a wide amount of simultaneous resources requests.

The users are localized at the cell edge, so there is an high amount of needed PRBs per user. We will show the results of the simulations, to evaluate how these conditions change the performances of the system. All the simulation parameters are shown in table 5.8.

Table 5.8: Simulation parameters Sim #5.8

Simulation parameter	Value
Simulation Time	80 s
Users in the system	250
CQI/SRS selectivity	Full CQI
CQI/SRS Reporting	10ms
DRX/DTX Cycle	$[40, 60, 80, 160]$ms
short DRX/DTX cycle	10ms
on-duration	$[5, 7]$ms

Figure 5.11 - The power consumption results are generally congruent to those of single user scenario. The average of results is however slightly higher, and this can be explained by the fact that the scheduler does not generally allocate the best PRBs for each UE, maybe because the best ones indicated by UE are already reserved for other users. The utilization of lower quality PRBs forces an higher BLEP that triggers more power consuming for retransmissions.

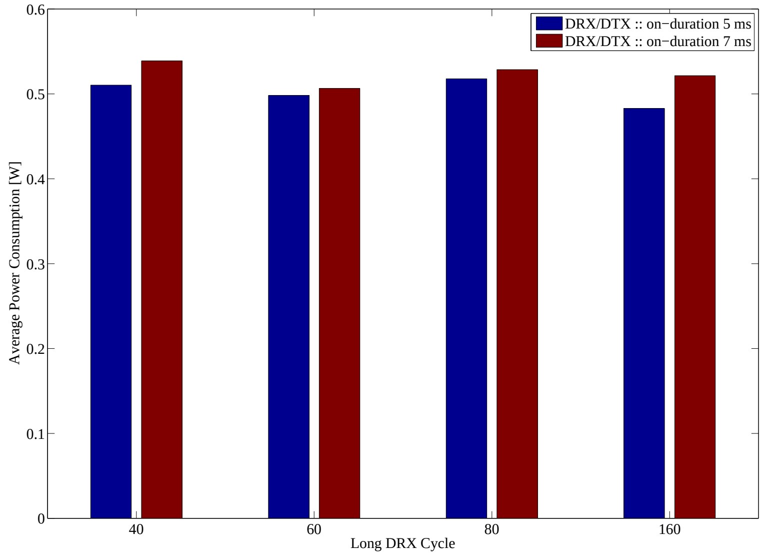


Figure 5.11: DS - Average power consumption, shortDRX, CQI/SRS periodicity 10ms, multiple users scenario (250 users at cell edge)

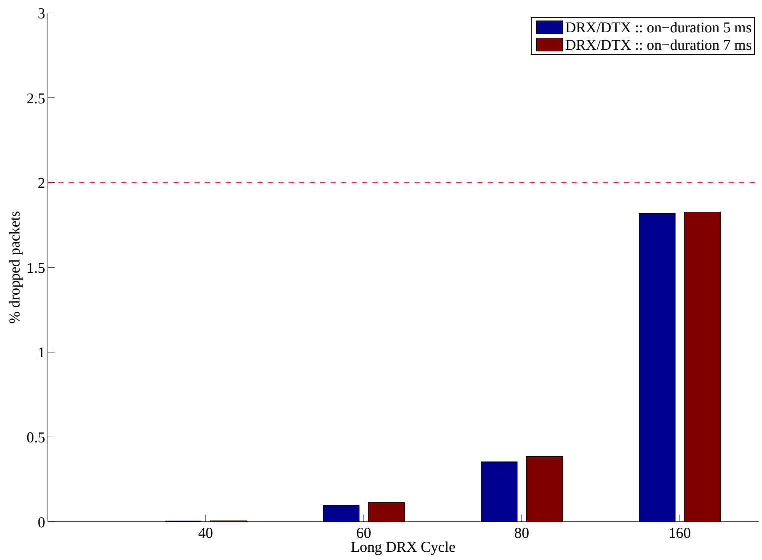


Figure 5.12: DS - Downlink packet drop probability, shortDRX, CQI/SRS periodicity 10ms, multiple users scenario (250 users at cell edge)

Figure 5.12 - The downlink packet drop probability is similar to the one of single user scenario. The expected degradation in scheduling performances is not so considerable. This is caused by the provided parameters that keep the UE in active state for a very big part of the DRX/DTX cycle: setting the short DRX/DTX cycle to 10ms and on-duration to 5 or 7 ms, the UE goes to sleep state only for a 50% or 30% of the time during a talk spurt. This makes space to an high number of eligible PRBs to perform the allocation.

The uplink delays results show that each packet does not suffer of any type of delay. Even the multiply retransmitted packets keep the transmission delay under 50 ms.

5.4.3 Single user scenario, CQI reporting period 20ms

The following simulations involve a larger CQI/SRS reporting period synchronized with the AMR codec periodicity. The first consequence of this choice is a lower power consumption, assuming that the transmissions of channel quality feedback are uplink transmissions, those require usually high power (figure 3.35). The settings concerning long DRX/DTX cycle and on-duration timers are the same of the previous analysis.

Table 5.9: Simulation parameters Sim #5.9

Simulation parameter	Value
Simulation Time	80 s
Users in the system	1
CQI/SRS selectivity	Full CQI
CQI/SRS Reporting	20ms
DRX/DTX Cycle	$[40, 60, 80, 160]$ms
short DRX/DTX cycle	20ms
on-duration	$[5, 7]$ms

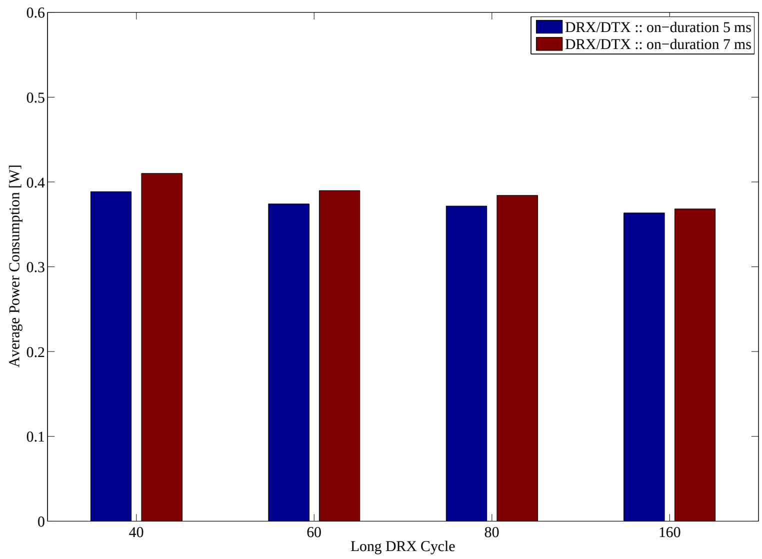


Figure 5.13: DS - Average power consumption, shortDRX, CQI/SRS periodicity 20ms, single user scenario

Figure 5.13 - The power consumption trend is quite similar to the one with 10 ms of short DRX/DTX cycle, however the average value for all the four sets of parameters is lower. This demonstrates that the CQI/SRS reporting activity, that runs with uplink transmissions, represent a strong component of the whole power consumption. The impact of this parameters change is higher than that

of long DRX/DTX cycle expansion, because of the general longer time the UE spends in short DRX/DTX condition.

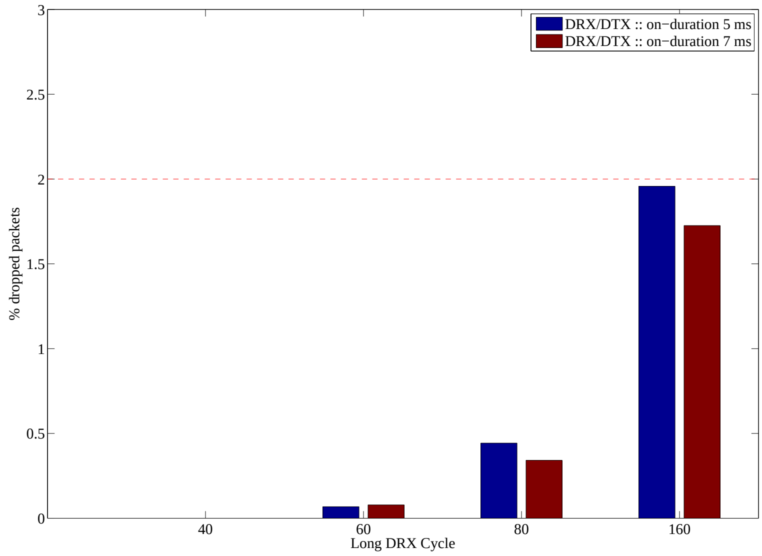


Figure 5.14: DS - Downlink packet drop probability, shortDRX, CQI/SRS periodicity 20ms, single user scenario

Figure 5.14 - The downlink drop probability for 20ms short DRX/DTX cycle follows the same trend of the 10 ms one. The explanation of the higher value of drop probability for 160 ms is well explained in section 5.4.1. An extended period for channel quality reporting does not show significant improvements in the overall user satisfaction.

Figure 5.15 - The uplink transmission does not show packets dropping events at all. The asynchronous character of uplink process makes near impossible an event like this with a single user in the system.

5.4.4 Multiple users scenario, CQI reporting period 20ms

Table 5.10: Simulation parameters Sim #5.10

Simulation parameter	Value
Simulation Time	80 s
Users in the system	250
CQI/SRS selectivity	Full CQI
CQI/SRS Reporting	20ms
DRX/DTX Cycle	$[40, 60, 80, 160]$ms
short DRX/DTX cycle	20ms
on-duration	$[5, 7]$ms

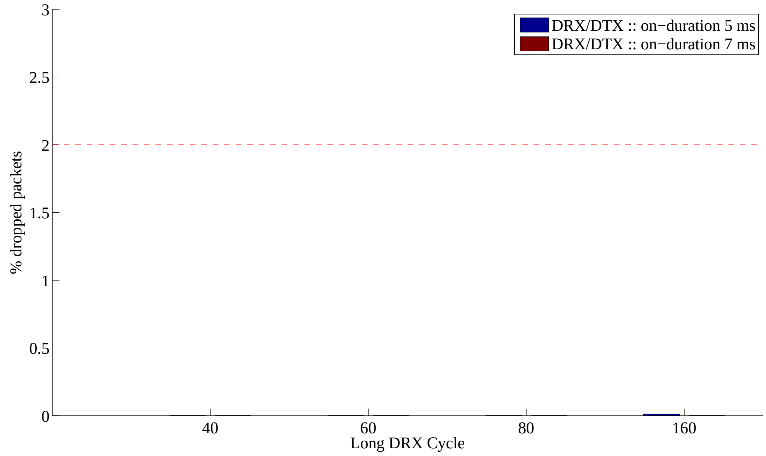


Figure 5.15: DS - Uplink packet drop probability, CQI/SRS periodicity 20ms, single user scenario

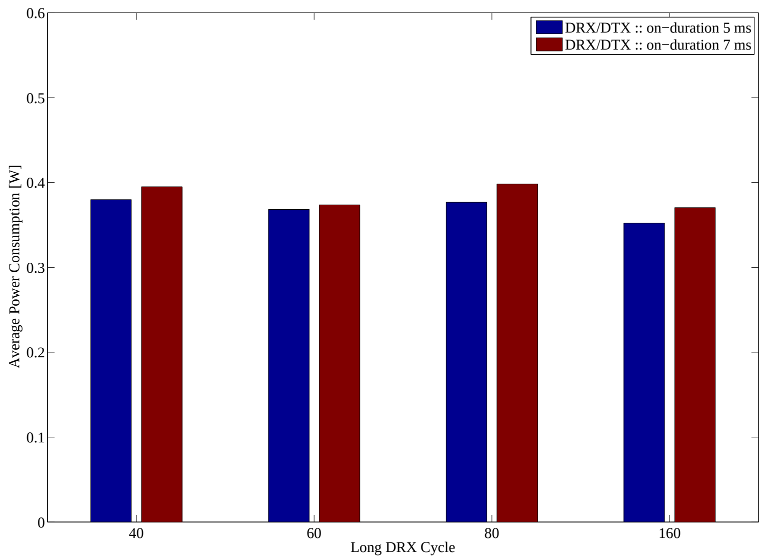


Figure 5.16: DS - Average power consumption, shortDRX, CQI/SRS periodicity 20ms, multiple users scenario (250 users at cell edge)

Figure 5.16 - The average power consumption with a short DRX/DTX cycle of 20 ms is obviously lower than the case with 10ms. The obtained gains are quite evident, and put the question if the decay in user satisfaction can be acceptable with this configuration.

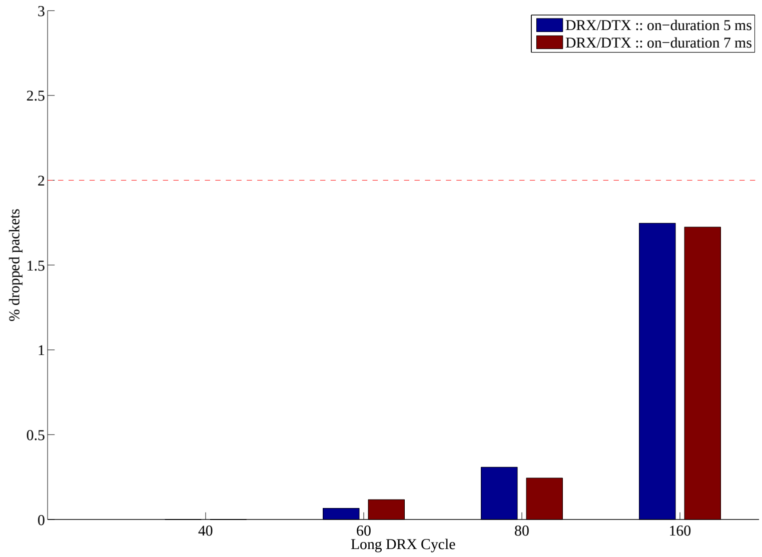


Figure 5.17: DS - Downlink packet drop probability, shortDRX, CQI/SRS periodicity 20ms, multiple users scenario (250 users at cell edge)

Figure 5.17 - The answer to the previous question is shown in downlink packet drop probability results. The threshold of 2% outage is never reached even with 160ms of DRX/DTX long cycle. This is due to the OLLA control that is able to work smoothly with a 20 ms clock, limiting the number of retransmissions to low values at the cost of a slight high number of used PRBs. The drop events are necessarily imputable to the beginning of talk spurt truncations previously explained.

Also in this scenario the uplink results do not show any useful behavior, and all packets are transmitted in the 50 ms threshold.

5.5 Main outcome

The performed analysis shows interesting results with the utilization of dynamic scheduler in power saving aspect with VoIP traffic. First of all it's unavoidable to note that the uplink results show an overall good user experience due to the fact that the UE can wake up asynchronously when it has data to transmit, and the eNodeB can tune the eventual short DRX/DTX cycle to fit the first available TTI. A second noticeable result is shown in short DRX/DTX cycles utilization: the power saving gains that it can introduce are very limited, so they do not seem the right direction to search for a reduction of power consumptions.

Chapter 6

SPS/DS Comparison

In this chapter we will take the results of the previous analysis to make a direct comparison between semi-persistent and dynamic packet scheduling policies. Both of enabled and disabled short DRX/DTX cases will be shown to verify pro and cons of each of the two solutions. The main outcome we want to find is the solution that can guarantee a trade-off between acceptable delays and significant power saving. For this reason the previous results are shown in single plots relating average power consumption and dropped packets.

6.1 Single user scenario

The first comparison is about the downlink scheduling and power saving performances in an ideal case that involves a completely empty cell scenario with a single user at the cell edge.

The SPS shows in general a very remarkable power saving characteristic but at the same time can not ensure the fulfillment of delay requirements with extended long DRX/DTX cycle (figure 6.1).

On the contrary the DS is able to make the scheduling to satisfy the user experience without high degradation 6.2. The cost of this peculiarity is obviously observable in the power consumption increase: the higher amount of signaling impacts to the average consumption.

As concern the uplink, the values of delays are not significant both for SPS and DS (figures 6.3 and 6.4). This happens because the UE can instantly require resources when it detects a new packet in its transmission buffer whatsoever the DRX/DTX state it is in. Moreover all the network resources are free in this scenario.

These first results show the impossibility to reach a compromise between power consumption saving and packets delays both for SPS and DS: if the SPS can assure more saving gain, at the same time its combination with the DRX/DTX functionalities decreases the SPS performance in packets delivering delays. At the contrary, the DS can guarantee lower delays but with more power consumption.

If we consider also the BLEP performance (section 4.2 and 5.2), the SPS with DRX/DTX can assure lower power consumption but with higher delays and an higher number of retransmissions due to the low precision to track the

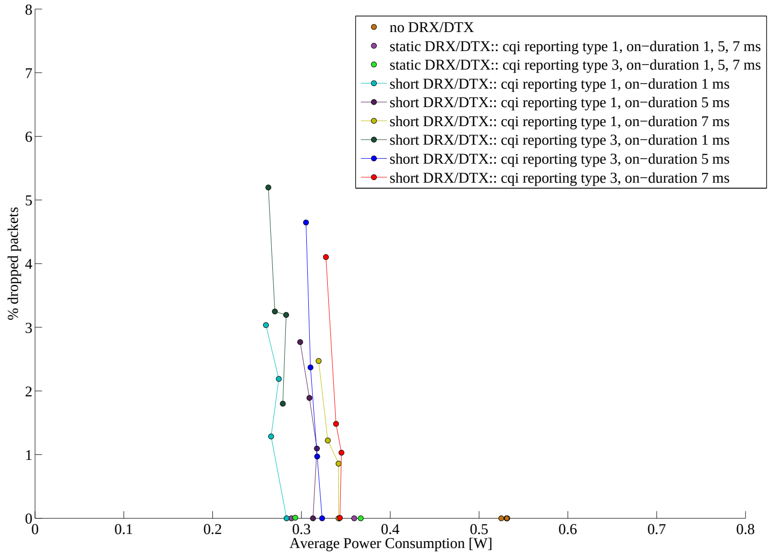


Figure 6.1: SPS - Average power consumption versus downlink packet drop probability, single user scenario

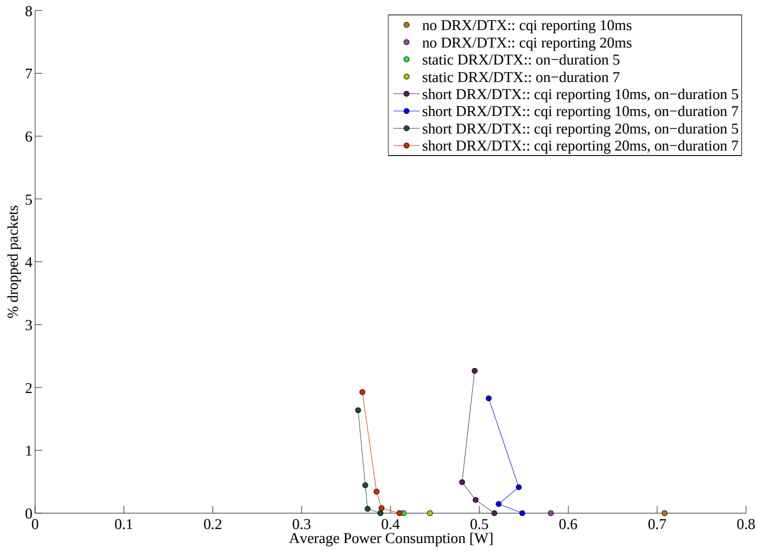


Figure 6.2: DS - Average power consumption versus downlink packet drop probability, single user scenario

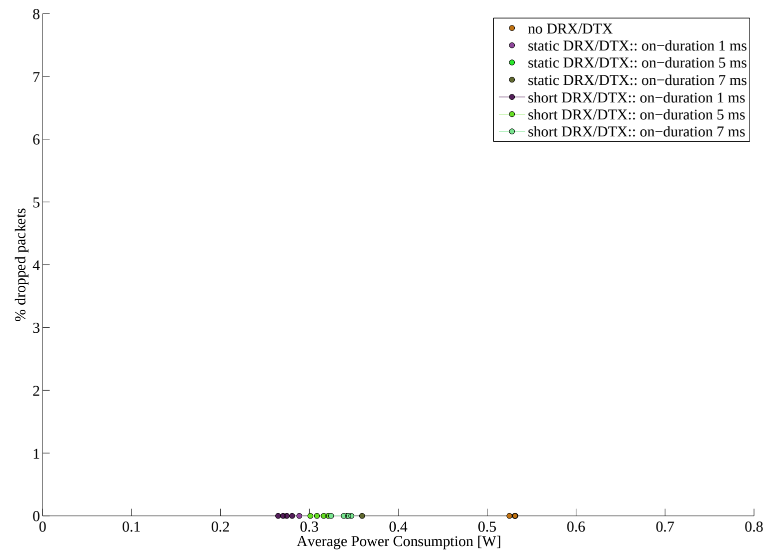


Figure 6.3: SPS - Average power consumption versus uplink packet drop probability, single user scenario

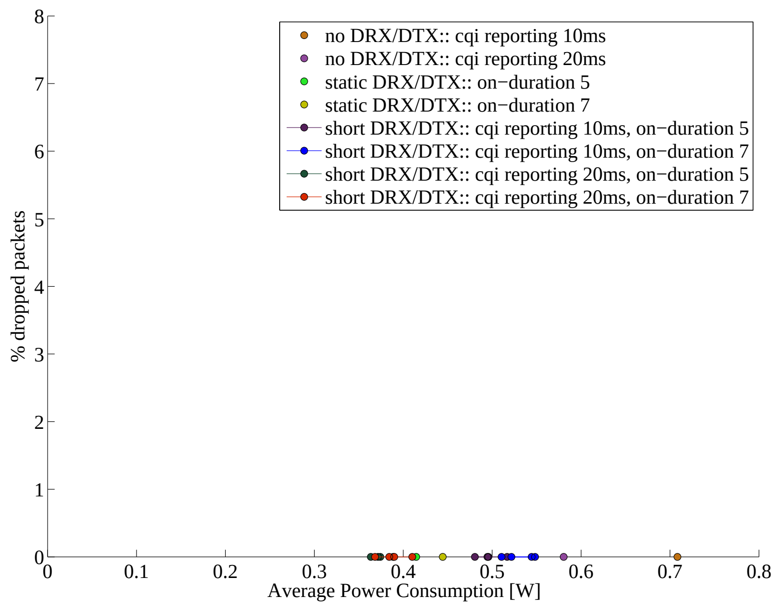


Figure 6.4: DS - Average power consumption versus uplink packet drop probability, single user scenario

BLEP target, especially in uplink. The DS, indeed, can fit the channel condition variations providing the BLEP target but with less power saving because of an increased control signaling.

6.2 Multiple users scenario

Varying the number of users in the cell, the network resources are obviously reduced because they are shared among an higher number of users, all of them at the cell edge. In these conditions, since there are fewer available PRBs per TTI, the scheduler has to search further in time for a possible allocation, increasing the packets delays (figures 6.5, 6.6, 6.7 and 6.8). This is worsened by the fact that we do not consider and implement VoIP packets segmentation and by extra constraints set by DRX/DTX activity.

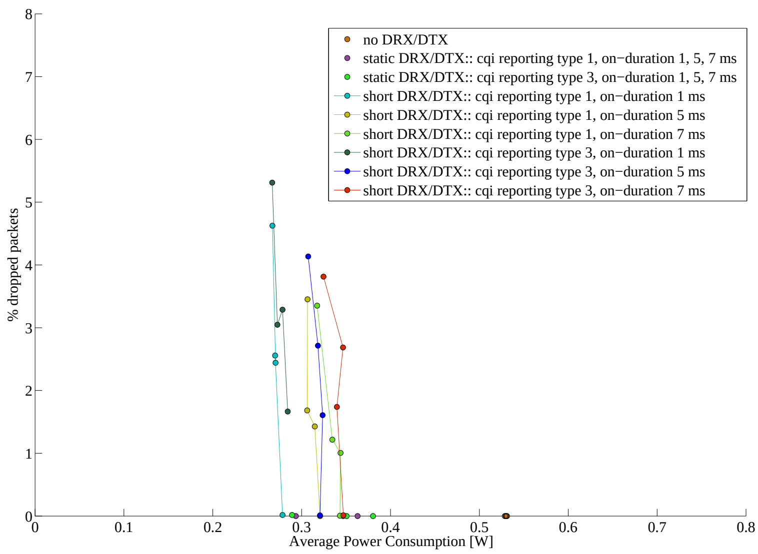


Figure 6.5: SPS - Average power consumption versus downlink packet drop probability, multiple users scenario (250 users)

In multiple users scenario, the trade-off between power saving and user satisfaction seems to be further compromised than the single one, both of SPS and DS. In particular, the delays are increased and we can suppose that with a number of users higher than 250 the performance can be further worsened. The power consumption, indeed, in not affected by the number of users in the cell.

6.3 Main outcome

The purpose of this chapter has been to summarize the results of the previous chapters 4 and 5, making a comparison between SPS and DS performance with DRX/DTX and trying to find a trade-off between power saving and user satisfaction.

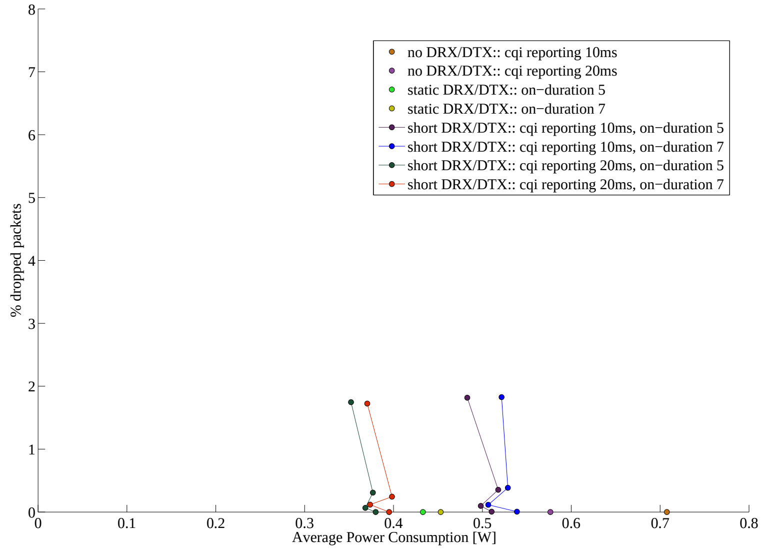


Figure 6.6: DS - Average power consumption versus downlink packet drop probability, multiple users scenario (250 users)

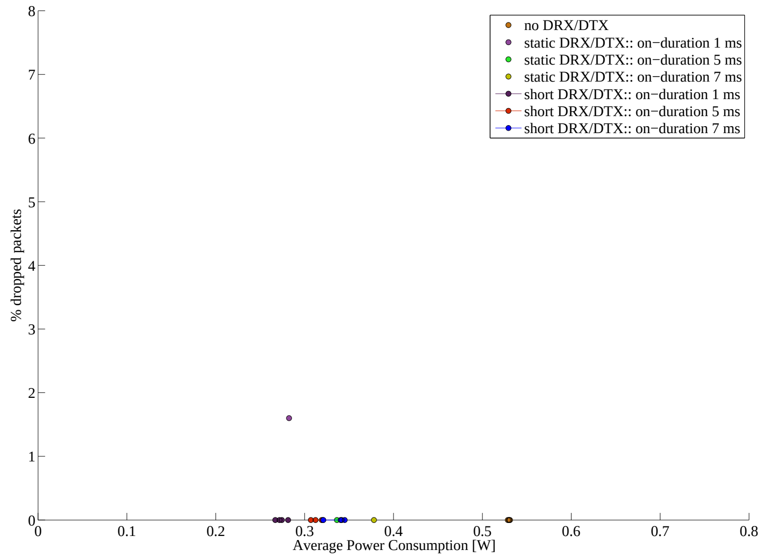


Figure 6.7: SPS - Average power consumption versus uplink packet drop probability, multiple users scenario (250 users)

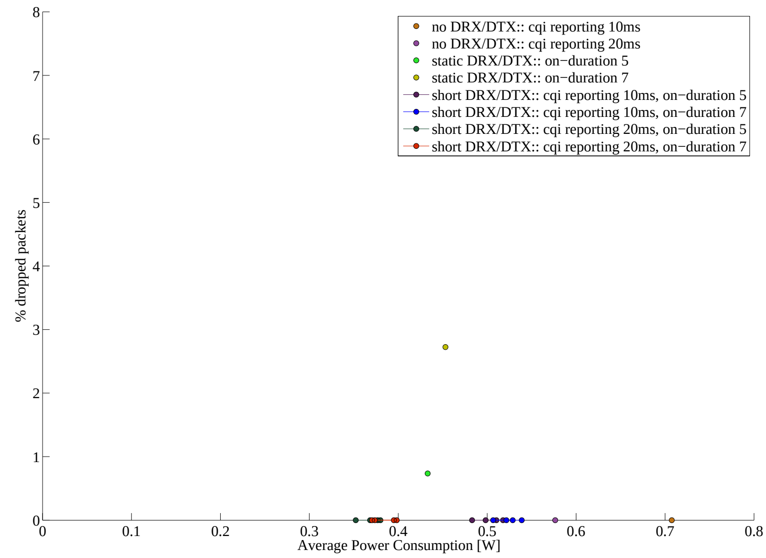


Figure 6.8: DS - Average power consumption versus uplink packet drop probability, multiple users scenario (250 users)

As concern the performance, the SPS and DS have opposing behaviour handling the UE with DRX/DTX: while the former shows a low UE's power consumption and high delays (especially in multi users scenario), the latter can provide a reliable transmission with lower delays but an higher power consumption for the UE. This means that a more frequent signaling can affect consistently the battery life, since the uplink transmissions imply high power consumption.

The trade-off between power saving and user satisfaction generally is an hard result to find. The best performances are shown by the SPS with long DRX of 40 ms and on-duration of 1 ms.

The overall analysis shows that the introduction of DRX/DTX feature, short DRX/DTX enabled or disabled, provides a reduction of power consumption values, and, if not correctly configured, can allow a little degradation of performances that is still not perceptible from the point of view of the user. The short DRX/DTX does not provide significant power savings, both for SPS and DS; the use of this functionality seems not to be worthy.

Chapter 7

Conclusions

The purpose of this master thesis project has been to evaluate the DRX/DTX functionality performances under semi-persistent and dynamic packet scheduling policies in order to find the best solution that can provide significant battery power saving for an LTE UE engaged in a VoIP call. The main chosen satisfaction indicators have been the average power consumption of the RF modem, and the packets delivery delays and dropping probability. The former gives an idea of the activity of the user equipment, the latter evaluates the actual satisfaction and usage experience for the user engaged in the VoIP call.

The project line of development involved the accurate modeling and implementation of a conversational speech model, the realization of the main LTE functionalities concerning scheduling and data transmission with link adaptation techniques, the creation of a specialized DRX/DTX framework and finally the implementation of a reference power consumption model.

An extensive simulation campaign has been carried out varying a wide number of parameter configurations, to be able to obtain a large results data base. This helped to find the best scheduling and DRX/DTX parameters as well as to verify the common behaviors of performances indicators under different assumptions and conditions. The tuning of inputs and parameters followed the specifications established by 3GPP and personal assumption derived from observations, analysis and discussions.

The SPS set of simulations emphasized the common defects that the *a priori* scheduling implies: the reduced amount of signaling costs an incorrect estimation of the effectively needed resources. Although the effects of SPS can result in both overestimation and underestimation of the channel quality, the observed trend is the former: without a proper control and preprocessing of inputs, the scheduler tends to assign too few resource units to the UE, resulting in an unreliable transmission. This is more evident in uplink (figure 3.25). The implemented controls named time aggressiveness and frequency aggressiveness, show that the SPS scheduler can be directed to better evaluate the channel conditions with a very simple calculation, leading the scheduler to commonly assign the same average number of PRBs (3.5 downlink, 7.5 uplink) that Dynamic scheduler chooses with the help of a more refined (and signaling expensive) control, as OLLA (figures 3.28, 3.13 and 3.14).

The increased robustness of transmissions obtained with the scheduler preprocessing control seems to work very well with downlink (figure 3.19), but it suffers of the high variance (±10%) of uplink channel quality, being able to assure a BLEP target of only 15% (figure 3.23). Although the number of SPS allocated PRBs is similar to the one of DS, the causes of higher uplink BLEP trend reside in the actual positioning of allocated PRBs. The preallocation and adjacency of resource units suffers of eventual bad positioning along the bandwidth and the impossibility to exploit frequency diversity.

The channel aware dynamic scheduling, run in the second set of simulations, assures an almost continue correctness in resources allocations. The downlink analysis, as well as in SPS, shows an accurate use of resources even without particular controls on the scheduler behavior (figure 3.13). In uplink, the OLLA action seems to be mandatory (figure 3.14). The OLLA control is demonstrated to be a good solution that shows its benefits providing accuracy regardless the channel quality feedback periodicity (figures 3.11 and 3.12).

The DRX/DTX functionality is an effective way to reduce the UE's battery power usage, but at the same time introduces further constraints in the scheduler's tasks. The immediate consequence of them is an average increase of packets delivery delays . The short DRX/DTX represents a further attempt to exploit the inactivity periods of UE to save even more power. This further saving could be remarkable with certain types of traffic, but can also be very limited with others, like VoIP.

DRX/DTX and SPS show a possible interoperability that can finely handle the particular type of network traffic represented by the VoIP. The known characteristics of this data source fits the SPS recurrent resource allocation that is adapted to the packets inter arrival time. The DRX/DTX can also overlap with this periodical activity realizing a system almost fully synchronized in all its parts. This is shown in the average good delay results of the SPS scheduler even for heavy loaded system (figures 4.8 and 4.9). This consideration, however, can not be fixed as a rule. If the number of users increases, the scheduler can have difficulties planning the allocation in the few time instants where the UE is supposed to be active and the packet scheduled, because of the lack of available resources. This is reflected in the waiting for a new UE's active state and a constant delay, propagated then on the whole talk-spurt (if TTI bundling is not implemented). While this behavior is very limited in static DRX/DTX simulations, it is clearly observable with the introduction of short DRX/DTX and high extended long DRX/DTX cycles (til +5% of dropped packets in downlink, figures 4.17 and 4.17). The small power savings obtainable with short DRX/DTX causes an unbearable user dissatisfaction (figure 6.5).

The DS scheduler interaction with DRX/DTX functionality shows a low induction of packets delivery delay: the more careful usage of radio resources can limit in some cases the number of retransmissions, and, moreover, the absence of preallocation avoids the propagation of delays along the talk-spurt. In fact each single packet transmission is not affected by previous transmission delays at all, unlike SPS. The downside of the use of this kind of scheduler is mainly represented by the amount of signaling activity that implies uplink transmission of CQI and SRS. This is worsened by the time between the request for channel reporting and transmission instant where UE remains in active state, affecting the battery life higher than SPS case (til +200 mW, figures 4.5 and 5.5 5.13). The increase of channel quality reporting frequency during talk-spurts (+50%) increases the amount of power consumption (figures 5.16 and 5.11), but does not improve the delays trend (figures 5.17 and 5.12).

One of the remark of this project concerns the not significant gain carried by short DRX/DTX functionality. The lack of big mutual silence period in a voice call makes the extension of DRX/DTX cycle really insignificant, in the attempt to save more power. Besides, the downlink responsiveness to new vocal activity is severely lowered, causing high delays and loss of a considerable amount of packets at the beginning of each talk spurt (from 3% to 6% of dropped packets).

7.1 Future works

The tool developed to perform this thesis analysis lends itself to be used to perform a quite big number of other studies that exploit any sort of combinations of the input parameters, added to the ones presented in the previous chapters. The system indeed can be fed with different link-level traces to test the system-level behavior for users in better conditions than at the cell edge. The behavior of all transmission entities and models can be finely tuned making very easy, for example, to apply other kinds of scheduling policies, DRX/DTX behaviors (battery-aware strategies, for example) or power consumption models.

It should be feasible with further analysis of scheduling behaviors, in case of a combined source of VoIP traffic and best effort traffic (web browsing, mail, FTP, etc.) too.

Bibliography

[1] 3GPP. Proposed study item on evolved utra and utran. RP 040461, 3rd Generation Partnership Project (3GPP), Athens, Greece, December 2004.

[2] 3GPP. Deployment aspects. TR 25943, 3rd Generation Partnership Project (3GPP), December 2008.

[3] 3GPP. Performance characterization of the adaptive multi-rate (amr) speech codec. Tr 26975, 3rd Generation Partnership Project (3GPP), dec 2008.

[4] 3GPP. Requirements for evolved utra (e-utra) and evolved utran (e-utran). TR 25913, 3rd Generation Partnership Project (3GPP), December 2008.

[5] 3GPP. (e-utra); medium access control (mac) protocol specification. Ts 36321, 3rd Generation Partnership Project (3GPP), mar 2009.

[6] 3GPP. Evolved universal terrestrial radio access (e-utra); physical layer procedures. Ts 36213, 3rd Generation Partnership Project (3GPP), may 2009.

[7] 3GPP. Radio resource control (rrc); protocol specification. Ts 36331, 3rd Generation Partnership Project (3GPP), jun 2009.

[8] Agilent. 3gpp long term evolution - system overview, product development, and test challenges. Technical report, Agilent Technologies, 2009.

[9] Syed Ahson. *Voip Handbook*. CRC, Boca Raton, 2008.

[10] Paul T. Brady. A model for generating on-off speech patterns in two-way conversation. *Bell System Technical Journal*, 1:2445–2472, 1969.

[11] Ericsson. Lte - an introduction. Technical Report 284 23-3124, Ericsson, jun 2009.

[12] Kolding T. E. Frederiksen, F. Performance and modeling of wcdma/hsdpa transmission/h-arq schemes. In *Proc. VTC 2002-Fall Vehicular Technology Conference 2002 IEEE 56th*, volume 1, pages 472–476 vol.1, 2002.

[13] Harri Holma. *Lte for Umts - Ofdma and Sc-Fdma Based Radio Access*. Wiley & Sons, New York, 2009.

[14] J. M. Holtzman. Asymptotic analysis of proportional fair algorithm. In *Proc. 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 2, pages F–33–F–37 vol.2, 2001.

[15] ITU-T. ITU-T recommendation G.114. Technical report, International Telecommunication Union, 1993.

[16] Farooq Khan. *Lte for 4g Mobile Broadband.* Cambridge University Press, Cambridge, 2009.

[17] T. Kolding, J. Wigard, and L. Dalsgaard. Balancing power saving and single user experience with discontinuous reception in lte. In *Proc. Wireless Communication Systems. 2008. ISWCS '08. IEEE International Symposium on*, pages 713–717, 2008.

[18] Suk-Bok Lee, I. Pefkianakis, A. Meyerson, Shugong Xu, and Songwu Lu. Proportional fair frequency-domain packet scheduling for 3gpp lte uplink. In *Proc. IEEE INFOCOM 2009*, pages 2611–2615, 2009.

[19] P. Lunden and M. Kuusela. Enhancing performance of voip over hsdpa. In *Proc. VTC2007-Spring Vehicular Technology Conference IEEE 65th*, pages 825–829, April 22–25, 2007.

[20] Emma Minoli and Daniel Minoli. *Delivering Voice over Ip Networks.* John Wiley & Sons, New York, 1998.

[21] G. Mongha, K. I. Pedersen, I. Z. Kovacs, and P. E. Mogensen. Qos oriented time and frequency domain packet schedulers for the utran long term evolution. In *Proc. IEEE Vehicular Technology Conference VTC Spring 2008*, pages 2532–2536, 2008.

[22] Nokia. Drx parameters in lte. R2 071285, 3rd Generation Partnership Project (3GPP), 2007.

[23] Per Henrik Michaelsen NSN. Lte amore requirement specification - drx/dtx and energy consumption. Internal specification.

[24] Orange, China Mobile, KPN, NTT DoCoMo, Sprint, T-Mobile, Vodafone, and Telecom Italia. Lte physical layer framework for performance verification. R1 070674, 3rd Generation Partnership Project (3GPP), 2007.

[25] K. I. Pedersen, F. Frederiksen, T. E. Kolding, T. F. Lootsma, and P. E. Mogensen. Performance of high-speed downlink packet access in coexistence with dedicated channels. 56(3):1262–1271, 2007.

[26] A. Pokhariyal, T. E. Kolding, and P. E. Mogensen. Performance of downlink frequency domain packet scheduling for the utran long term evolution. In *Proc. IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–5, September 11–14, 2006.

[27] Stefania Sesia. *Lte, the Umts Long Term Evolution.* Wiley, New York, 2009.

[28] Guocong Song and Ye Li. Utility-based resource allocation and scheduling in ofdm-based wireless broadband networks. 43(12):127–134, 2005.

[29] Jacob Sorber, Nilanjan Banerjee, Mark D. Corner, and Sami Rollins. Turducken: hierarchical power management for mobile devices. In *MobiSys '05: Proceedings of the 3rd international conference on Mobile systems, applications, and services*, pages 261–274, New York, NY, USA, 2005. ACM.

[30] Lin Zhong and Niraj K. Jha. Energy efficiency of handheld computer interfaces: limits, characterization and practice. In *MobiSys '05: Proceedings of the 3rd international conference on Mobile systems, applications, and services*, pages 247–260, New York, NY, USA, 2005. ACM.

List of Figures

List of Tables